FIGURE 1

GTTACTCGGTGGTGGCGGAGTCTACGGAAGCCGTTTTCGCTTCACTTTTCCTGGCTGTAGAGC

GCTTTCCCCCTGGCGGGTGAGAGTGCAGAGACGAAGGTGCGAG**ATG**AGCACTATGTTCGCGGA

CACTCTCCTCATCGTTTTTATCTCTGTGTGCACGGCTCTGCTCGCAGAGGGCATAACCTGGGT

CCTGGTTTACAGGACAGACAAGTACAAGAGACTGAAGGCAGAAGTGGAAAAACAGAGTAAAAA

ATTGGAAAAGAAGAAGGAAACAATAACAGAGTCAGCTGGTCGACAACAGAAAAGAAAATAGA

GAGACAAGAAGAGAAACTGAAGAATAACAACAGAGATCTATCAATGGTTCGAATGAAATCCAT

GTTTGCTATTGGCTTTTGTTTTACTGCCCTAATGGGAATGTTCAATTCCATATTTGATGGTAG

AGTGGTGGCAAAGCTTCCTTTTACCCCTCTTTCTTACATCCAAGGACTGTCTCATCGAAATCT

GCTGGGAGATGACACCACAGACTGTTCCTTCATTTTCCTGTATATTCTCTGTACTATGTCGAT

TCGACAGAACATTCAGAAGATTCTCGGCCTTGCCCCTTCACGAGCCGCCACCAAGCAGGCAGG

TGGATTTCTTGGCCCACCACCTCCTTCTGGGAAGTTCTCT**TGA**ACTCAAGAACTCTTTATTTT

CTATCATTCTTTCTAGACACACACACATCAGACTGGCAACTGTTTTGTAGCAAGAGCCATAGG

TAGCCTTACTACTTGGGCCTCTTTCTAGTTTTGAATTATTTCTAAGCCTTTTGGGTATGATTA

GAGTGAAAATGGCAGCCAGCAAACTTGATAGTGCTTTTGGTCCTAGATGATTTTTATCAAATA

AGTGGATTGATTAGTTAAGTTCAGGTAATGTTTATGTAATGAAAAACAAATAGCATCCTTCTT

GTTTCATTTACATAAGTATTTTCTGTGGGACCGACTCTCAAGGCACTGTGTATGCCCTGCAAG

TTGGCTGTCTATGAGCATTTAGAGATTTAGAAGAAAAATTTAGTTTGTTTAACCCTTGTAACT

GTTTGTTTTGTTGTTGTTTTTTTTTCAAGCCAAATACATGACATAAGATCAATAAAGAGGCCA

AATTTTTAGCTGTTTTATGTACAAGGAGAGATCTGTTTCATTTTGTTTTGCCGTATTTCTAGA

TATAAGTTTTAGCATGGGCCAGGAAGGACTAAAATAAAGTTTTTAAGGTACAAAAAAAAAAA

AAAA

FIGURE 2

MSTMFADTLLIVFISVCTALLAEGITWVLVYRTDKYKRLKAEVEKQSKKLEKKKETITESAGR

QQKKKIERQEEKLKNNNRDLSMVRMKSMFAIGFCFTALMGMFNSIFDGRVVAKLPFTPLSYIQ

GLSHRNLLGDDTTDCSFIFLYILCTMSIRQNIQKILGLAPSRAATKQAGGFLGPPPPSGKFS

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 103-109, 163-169

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 53-57

FIGURE 3

AGCCGGGGGCGGGTTTGAAGACGCGTCGTTGGGTTTTGGAGGCCGTGAAACAGCCGTTTGAGT

TTGGCTGCGGGTGGAGAACGTTTGTCAGGGGCCCGGCCAAGAAGGAGGCCCGCCTGTTACG**AT**

GGTGTCCATGAGTTTCAAGCGGAACCGCAGTGACCGGTTCTACAGCACCCGGTGCTGCGGCTG

TTGCCATGTCCGCACCGGGACGATCATCCTGGGGACCTGGTACATGGTAGTAAACCTATTGAT

GGCAATTTTGCTGACTGTGGAAGTGACTCATCCAAACTCCATGCCAGCTGTCAACATTCAGTA

TGAAGTCATCGGTAATTACTATTCGTCTGAGAGAATGGCTGATAATGCCTGTGTTCTTTTTGC

CGTCTCTGTTCTTATGTTTATAATCAGTTCAATGCTGGTTTATGGAGCAATTTCTTATCAAGT

GGGTTGGCTGATTCCATTCTTCTGTTACCGACTTTTTGACTTCGTCCTCAGTTGCCTGGTTGC

TATTAGTTCTCTCACCTATTTGCCAAGAATCAAAGAATATCTGGATCAACTACCTGATTTTCC

CTACAAAGATGACCTCCTGGCCTTGGACTCCAGCTGCCTCCTGTTCATTGTTCTTGTGTTCTT

TGCCTTATTCATCATTTTTAAGGCTTATCTAATTAACTGTGTTTGGAACTGCTATAAATACAT

CAACAACCGAAACGTGCCGGAGATTGCTGTGTACCCTGCCTTTGAAAGCACCTCCTCAGTACG

TTTTGCCAACCTATGAAATGGCCGTGAAAATGCCTGAAAAAGAACCACCACCTCCTTACTTAC

CTGCCTGAAGAAATTCTGCCTTTGACAATAAATCCTATACCAGCTTTTTGTTTGTTTATGTTA

CAGAATGCTGCAATTCAGGGCTCTTCAAACTTGTTTGATATAAAATATGTTGTCTTTTGTTTA

AGCATTTATTTTCAAACACTAAGGAGCTTTTTGACATCTGTTAAACGTCTTTTTGTTTTTTTG

TTAAGTCTTTTACATTTTAATAGTTTTTGAAGACAATCTAGGTTAAGCAAGAGCAAAGTGCCA

TTGTTTGCCTTTAATTGGGGGGTGGGAAGGGAAAGAGGGTACTTGCCACATAGTTTCCTTTTT

AACTGCACTTTCTTTATATAATCGTTTGCATTTTGTTACTTGCTACCCTGAGTACTTTCAGGA

AGACTGACTTAAATATTCGGGGTGAGTAAGTAGTTGGGTATAAGATCTGAACTTTTCATCTGC

AGAGGCAAGAAAAATATTTGACATTGTGACTTGACTGTGGAAGATGATGGTTGCATGTTTCTA

GTTTGTATATGTTTCCATCTTTGTGATAAGATGATTTAATAAATCTCTTTAAATACTAAAAAA

AAAAAAAAA

FIGURE 4

MVSMSFKRNRSDRFYSTRCCGCCHVRTGTIILGTWYMVVNLLMAILLTVEVTHPNSMPAVNIQ

YEVIGNYYSSERMADNACVLFAVSVLMFIISSMLVYGAISYQVGWLIPFFCYRLFDFVLSCLV

AISSLTYLPRIKEYLDQLPDFPYKDDLLALDSSCLLFIVLVFFALFIIFKAYLINCVWNCYKY

INNRNVPEIAVYPAFESTSSVRFANL

Important features of the protein:

Transmembrane domain (Possible type II transmembrane protein):

amino acids 30-49, 81-100, 111-131, 158-175

N-glycosylation site.

amino acids 9-13

Tyrosine kinase phosphorylation sites.

amino acids 8-16, 193-202

N-myristoylation site.

amino acids 68-74

FIGURE 5

CCCGCTGGCCCGTCAGTGCTCTCCCCGTCGTTTGCCCTCTCCAGTTCCCCCAGTGCCTGCCCT

ACGCACCCCG**ATG**GCGGAGCTGCGGCCTAGCGGCGCCCCCGGCCCCACCGCGCCCCGGCCCC

TGGCCCGACTGCCCCCCCGGCCTTCGCTTCGCTCTTTCCCCCGGGACTGCACGCCATCTACGG

AGAGTGCCGCCGCCTTTACCCTGACCAGCCGAACCCGCTCCAGGTTACCGCTATCGTCAAGTA

CTGGTTGGGTGGCCCAGACCCCTTGGACTATGTTAGCATGTACAGGAATGTGGGGAGCCCTTC

TGCTAACATCCCCGAGCACTGGCACTACATCAGCTTCGGCCTGAGTGATCTCTATGGTGACAA

CAGAGTCCATGAGTTTACAGGAACAGATGGACCTAGTGGTTTTGGCTTTGAGTTGACCTTTCG

TCTGAAGAGAGAAACTGGGGAGTCTGCCCCACCAACATGGCCCGCAGAGTTAATGCAGGGCTT

GGCACGATACGTGTTCCAGTCAGAGAACACCTTCTGCAGTGGGGACCATGTGTCCTGGCACAG

CCCTTTGGATAACAGTGAGTCAAGAATTCAGCACATGCTGCTGACAGAGGACCCACAGATGCA

GCCCGTGCAGACACCCTTTGGGGTAGTTACCTTCCTCCAGATCGTTGGTGTCTGCACTGAAGA

GCTACACTCAGCCCAGCAGTGGAACGGGCAGGGCATCCTGGAGCTGCTGCGGACAGTGCCTAT

TGCTGGCGGCCCCTGGCTGATAACTGACATGCGGAGGGGAGAGACCATATTTGAGATCGATCC

ACACCTGCAAGAGAGAGTTGACAAAGGCATCGAGACAGATGGCTCCAACCTGAGTGGTGTCAG

TGCCAAGTGTGCCTGGGATGACCTGAGCCGGCCCCCCGAGGATGACGAGGACAGCCGGAGCAT

CTGCATCGGCACACAGCCCCGGCGACTCTCTGGCAAAGACACAGAGCAGATCCGGGAGACCCT

GAGGAGAGGACTCGAGATCAACAGCAAACCTGTCCTTCCACCAATCAACCCTCAGCGGCAGAA

TGGCCTCGCCCACGACCGGGCCCCGAGCCGCAAAGACAGCCTGGAAAGTGACAGCTCCACGGC

CATCATTCCCCATGAGCTGATTCGCACGCGGCAGCTTGAGAGCGTACATCTGAAATTCAACCA

GGAGTCCGGAGCCCTCATTCCTCTCTGCCTAAGGGGCAGGCTCCTGCATGGACGGCACTTTAC

ATATAAAGTATCACAGGTGACATGGCCATCACGTTTGTCTCCACGGGAGTGGAAGGCGCCTT

TGCCACTGAGGAGCATCCTTACGCGGCTCATGGACCCTGGTTACAACTC**TGA**ACCTATCCTCG

GAGCTCTGCCCTCCCGTCCTGGAACGTCTTTCTGCCCTGAGGAGAGGGTAGTCAGCATCTCCA

ATTTTCAGCAGCTCAAGAACCTTGGCCCCCACAGGACTTCGCAGATGTCACATTGCCCCTCAG

TCCCCTGAATGCCCTTCGGACCCAACCCCAATTCCCCAAGCCCCTGACCCCCTAGCTGCCGGG

GTTCCCACTCCCAGTGCCACAACCCCCTCACCTCCCCTGGCAGCCCCTCAGCGAGCCTGAGGC

CCAGCACCCGCTGGCTCCCCAGCACATGGTCCCCTCCCATGGGCTGTTGCCCAGGGAACCGGG

GCGCGGTGGGAACGAGCTGCTGGCCTCGGCATGTTTCAATAAAGTTGCTGTGCTGGGAG

FIGURE 6

MAELRPSGAPGPTAPPAPGPTAPPAFASLFPPGLHAIYGECRRLYPDQPNPLQVTAIVKYWLG

GPDPLDYVSMYRNVGSPSANIPEHWHYISFGLSDLYGDNRVHEFTGTDGPSGFGFELTFRLKR

ETGESAPPTWPAELMQGLARYVFQSENTFCSGDHVSWHSPLDNSESRIQHMLLTEDPQMQPVQ

TPFGVVTFLQIVGVCTEELHSAQQWNGQGILELLRTVPIAGGPWLITDMRRGETIFEIDPHLQ

ERVDKGIETDGSNLSGVSAKCAWDDLSRPPEDDEDSRSICIGTQPRRLSGKDTEQIRETLRRG

LEINSKPVLPPINPQRQNGLAHDRAPSRKDSLESDSSTAIIPHELIRTRQLESVHLKFNQESG

ALIPLCLRGRLLHGRHFTYKSITGDMAITFVSTGVEGAFATEEHPYAAHGPWLQL

Important features:

N-glycosylation site.

amino acids 265-268

FIGURE 7

CGCGAATGAAGTTTGCATTTTCCTCTGTTCTTGAGCCCAGCTTCTTCTCGTCTCCCACCCCAG

CTTCCCGGCATTGGAAGAAGGGACCGTCCTCTTCCTTGTCTTGGCCACCCAAATCCTGGTATC

GAAAGGGTTGAACGGACCGGAAGTGTGCAGCAGCGACGGGTCCCCAGCTAATCGACGCCGGAA

GTAGCAATTACTAGACAAGCATTCCGCCGCCGGCTTCGCT**ATG**GCGGCAATTCCCCCAGATTC

CTGGCAGCCACCCAACGTTTACTTGGAGACCAGCATGGGAATCATTGTGCTGGAGCTGTACTG

GAAGCATGCTCCAAAGACCTGTAAGAACTTTGCTGAGTTGGCTCGTCGAGGTTACTACAATGG

CACAAAATTCCACAGAATTATCAAAGACTTCATGATCCAAGGAGGTGACCCAACAGGGACAGG

TCGAGGTGGTGCATCTATCTATGGCAAACAATTTGAAGATGAACTTCATCCAGACTTGAAATT

CACGGGGGGCTGGAATTCTCGCAATGGCCAATGCGGGGCCAGATACCAATGGCAGCCAGTTCTT

TGTGACCCTCGCCCCCACCCAGTGGCTTGACGGCAAACACACCATTTTTGGCCGAGTGTGTCA

GGGCATAGGAATGGTGAATCGCGTGGGAATGGTAGAAACAAACTCCCAGGACCGCCCTGTGGA

CGACGTGAAGATCATTAAGGCATACCCTTCTGGG**TAG**ACTTGCTACCTCTTGAGCAGCTCTT

CTGAGATGGCCCCAGTGAACCAGCTTCTAGATGACATAGAATGACATGTAATGCTAAATTTCA

TTTTGGCTTTGCAAGTCATGAAGCTTAGGAGGCCTGGCATCTTGGGTGAGTTAGAGATGGAAG

TACATTTTAATAGGATGCTTCTTTTCTCTTCCCCCAGTGCCTAGGTTGCCAGAGCATTTGCAC

AAATGCCCCTGTTTATCAATAGGTGACTACTTACTACACATGAACCATAATGCTGCTTCTTGT

GCATGTCTGCTCTGATATACGTCGAACAATGTAGCAGCCACTGTCATTTCTCAGTGGTTTTGC

CTAACCAAACTTCTTCCTAAGGAGATTTATATTCTGGCCTACACAGCAGTCCTTGATGGCTGA

CAGCCACAGAATTCCAAACCAAGTAGTGTCTGTCAGCCCTCTTAACTCTGTGCACGCCCTATT

TCAGTCTTTTACATTTGTTCTTCTAGGGAATGTATGCATCTCTATATATATTTTCCCTCTCAA

AACCAGAACATCAACAGTGCTGTTTCTGACACTTCAGACATCCCACGCAAAGCCACATTGAAT

TTTTGCCAAATGAAAACACATCCAACAATCAAGTTTCTAAGAAGGTGTCAAGTGGGGAATAA

TAATAATGTATAATAATCAAGAAATTAGTTTATTAAAAGGAAGCAGAAGCATTGACCATTTTT

TCCCAGAGAAGAGGAGAAATCTGTAGTGAGCAAAGGACAGACCATGAATCCTCCTTGAGAAGT

AGTACTCTCAGAAAGGAGAAGCGCCACTCAAGTTCTTTTAACCCAAGACTTTAGAGAAATTAG

GTCCAAGATTTTTATATGTTCAGTTGTTTATGTATAAAAATAACTTTCTGGATTTTGTGGGGA

GGAGCAGGAGAGGAAGGAAGTTAATACCTATGTAATACATAGAAACTTCCACAATAAAATGCC

ATTGATGGTTAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MAAIPPDSWQPPNVYLETSMGIIVLELYWKHAPKTCKNFAELARRGYYNGTKFHRIIKDFMIQ

GGDPTGTGRGGASIYGKQFEDELHPDLKFTGAGILAMANAGPDTNGSQFFVTLAPTQWLDGKH

TIFGRVCQGIGMVNRVGMVETNSQDRPVDDVKIIKAYPSG

Important features:

N-glycosylation sites:

amino acids 49-52, 108-111

N-myristoylation sites:

amino acids 64-69, 69-74, 143-148

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature:

amino acids 48-65

FIGURE 9

```
CGGACGCGTGGGCGCGCGCGAGCGCAGCGGTGGGAGGCGGCGACCAGCCGGTTGAGGCCCCAG
GCTTGGCCTCACCACAATGTGGCACGAGGCTCGGAAGCATGAGCGGAAGCTTCGAGGCATGAT
GGTCGACTACAAGAAGAGGGCGGAGCGGAGACGGGAGTATTATGAAAAGATCAAGAAGGACCC
AGCCCAGTTCCTGCAGGTACATGGCCGAGCTTGCAAGGTGCACCTGGATTCTGCAGTCGCCCT
GGCCGCTGAGAGCCCTGTTAATATGATGCCCTGGCAGGGGGACACCAACAACATGATTGACCG
ATTCGATGTCCGTGCCCACCTGGACCACATCCCCGACTACACCCCCCCTCTGCTCACCACCAT
CTCCCCAGAACAGGAGTCGGACGAACGGAAGTGTAACTACGAGCGCTACAGAGGCCTGGTGCA
GAACGACTTTGCCGGCATCTCAGAGGAGCAGTGCCTGTACCAGATCTACATTGATGAGTTGTA
CGGAGGCCTCCAGAGACCCAGCGAAGATGAGAAGAAGAAGCTGGCAGAGAAGAAGGCTTCCAT
CGGTTATACCTACGAGGACAGCACGGTGGCCGAGGTAGAGAAGGCGGCAGAAAAGCCAGAGGA
GGAGGAGTCAGCGGCCGAGGAGGAGAGCAACTCGGACGAAGATGAGGTCATCCCCGACATCGA
CGTGGAGGTGGACGTGGATGAATTGAACCAGGAGCAGGTGGCAGATCTCAACAAACAGGCCAC
GACTTATGGCATGGCCGACGGTGACTTCGTCAGGATGCTCCGGAAAGACAAGGAGGAGGCAGA
GGCCATCAAGCATGCCAAGGCTCTTGAGGAGGAGAAGGCCATGTACTCGGGACGCCGCTCTCG
ACGCCAGCGGAGAGAGTTTCGGGAGAAGCGGCTGAGGGGTCGCAAGATCAGCCCACCCAGCTA
TGCCCGCCGAGACAGCCCCACCTATGACCCCTATAAGCGGTCACCCTCGGAGTCCAGCTCAGA
GTCCCGCTCCCGCTCCCGCTCCCCGACCCCGGGCCGCGAGGAGAAGATCACGTTCATCACCAG
TTTTGGGGGCAGCGATGAGGAGGCAGCCGCAGCCGCTGCTGCCGCAGCAGCATCAGGAGTCAC
CACAGGGAAGCCCCCCGCACCTCCCCAGCCTGGCGGCCCCGCCCCGGGACGTAATGCCAGCGC
CCGCCGCCGCTCCTCCTCCTCCTCCTCCTCCTCTTCTGCCTCGAGGACCTCCAGCTCCCGCTC
CAGCTCTCGCTCCAGCTCCCGCTCTCGCCGTGGTGGGGGCTACTACCGTTCCGGCCGCCACGC
CCGCTCCCGGTCCCGCTCCTGGTCCCGCTCCCGCTCCCGCTCCCGGCGCTATTCCCGGTCCCG
TAGCCGTGGCCGGCGGCACTCAGGTGGGGGCTCCCGAGACGGACACCGGTACTCCCGCTCGCC
CGCCCGGCGTGGTGGTTACGGGCCCCGGCGCAGAAGCAGGAGCCGCTCCCACTCAGGGGACCG
CTACAGGCGGGGCGGCCGGGGCCTCAGGCACCACAGCAGTAGCCGCAGCCGCAGCAGCTGGTC
CCTCAGCCCGTCCCGCAGTCGCAGCCTGACTCGCAGCCGCAGCCATAGCCCCAGCCCCAGCCA
GAGCCGCAGCCGCAGCCGCAGCCGCAGCCAGAGCCCCTCGCCATCACCCGCAAGAGAGAAGCT
GACCAGGCCGGCCGCGTCCCCTGCTGTGGGCGAGAAGCTGAAAAAGACCGAACCTGCCGCTGG
TAAAGAGACAGGAGCTGCCAAAGTCACCCAAGCTGACGCCTCAGGAGAAGCTGAAACTGAGGA
TGCAGAAGGCGCTGAACAGGCAGTTCAAGGCGGATAAGAAGGCGGCACAAGAAAAGATGATCC
AGCAGGAGCATGAGCGGCAGGAGCGGGAAGACGAGCTTCGAGCCATGGCCCGCAAGATCCGCA
TGAAGGAGCGGGAACGCCGAGAGAAGGAGAGAGAAGAGTGGGAACGCCAGTACAGCCGGCAGA
GCCGCTCACCCTCCCCCCGATACAGTCGAGAATACAGCTCTTCTCGAAGGCGCTCAAGGTCCC
GATCCCGAAGCCCCCATTACCGACATTAGGCAGAAGAGTGGGGGGTGGGGAGGACAAGGGGGT
GGGTAAGGGGCTCAAGCTGTGATGCTGCTGGTTTTATCTCTAGTGAAATAAAGTCAAAAGTTA
TTTAATTCCCGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAA
```

FIGURE 10

MWHEARKHERKLRGMMVDYKKRAERRREYYEKIKKDPAQFLQVHGRACKVHLDSAVALAAESP
VNMMPWQGDTNNMIDRFDVRAHLDHIPDYTPPLLTTISPEQESDERKCNYERYRGLVQNDFAG
ISEEQCLYQIYIDELYGGLQRPSEDEKKKLAEKKASIGYTYEDSTVAEVEKAAEKPEEEESAA
EEESNSDEDEVIPDIDVEVDVDELNQEQVADLNKQATTYGMADGDFVRMLRKDKEEAEAIKHA
KALEEEKAMYSGRRSRRQRREFREKRLRGRKISPPSYARRDSPTYDPYKRSPSESSSESRSRS
RSPTPGREEKITFITSFGGSDEEAAAAAAAAAASGVTTGKPPAPPQPGGPAPGRNASARRRSS
SSSSSSSASRTSSSRSSSRSSSRSRRGGGYYRSGRHARSRSRSWSRSRSRSRRYSRSRSRGRR
HSGGGSRDGHRYSRSPARRGGYGPRRRSRSRSHSGDRYRRGGRGLRHHSSSRSRSSWSLSPSR
SRSLTRSRSHSPSPSQSRSRSRSRSQSPSPSPAREKLTRPAASPAVGEKLKKTEPAAGKETGA
AKVTQADASGEAETEDAEGAEQAVQGG

Important features:

N-glycosylation site:

amino acids 370-373

Glycosaminoglycan attachment site:

amino acids 443-446

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 159-162, 282-285, 291-294, 374-377, 375-378, 430-433,
440-443, 466-469

Casein kinase II phosphorylation site:

amino acids 149-152, 166-169, 171-174, 187-190, 193-196, 195-198,
303-306, 307-310, 335-338, 571-574

N-myristoylation sites:

amino acids 118-123, 229-234, 350-355, 446-451, 586-591

Amidation sites:

amino acids 263-266, 280-283, 438-441

FIGURE 11

```
GGTAGGCGCGCCCAGACCTGAGACGGGTTGGGACTGGGCTGCGTCACGCGCGGGCTCTAAGCG
CCCGGGGCCCCGCCCAGTGGCCGGCACAGCCAATCGCAGCGCGGGAAGGCGGTGGGGGCGGGG
AAGGCCGCCTGGAAACTTAAATCCCGAGGCGGGCGAACCTGCACCAGACCGCGGACGTCTGTA
ATCTCAGAGGCTTGTTTGCTGAGGGTGCCTGCGCAGCTGCGACGGCTGCTGGTTTTGAAAC**AT**
**G**AATCTTTCGCTCGTCCTGGCTGCCTTTTGCTTGGGAATAGCCTCCGCTGTTCCAAAATTTGA
CCAAAATTTGGATACAAAGTGGTACCAGTGGAAGGCAACACACAGAAGATTATATGGCGCGAA
TGAAGAAGGATGGAGGAGAGCAGTGTGGGAAAAGAATATGAAAATGATTGAACTGCACAATGG
GGAATACAGCCAAGGGAAACATGGCTTCACAATGGCCATGAATGCTTTTGGTGACATGACCAA
TGAAGAATTCAGGCAGATGATGGGTTGCTTTCGAAACCAGAAATTCAGGAAGGGGAAAGTGTT
CCGTGAGCCTCTGTTTCTTGATCTTCCCAAATCTGTGGATTGGAGAAAGAAAGGCTACGTGAC
GCCAGTGAAGAATCAGAAACAGTGTGGTTCTTGTTGGGCTTTTAGTGCGACTGGTGCTCTTGA
AGGACAGATGTTCCGGAAAACTGGGAAACTTGTCTCACTGAGCGAGCAGAATCTGGTGGACTG
TTCGCGTCCTCAAGGCAATCAGGGCTGCAATGGTGGCTTCATGGCTAGGGCCTTCCAGTATGT
CAAGGAGAACGGAGGCCTGGACTCTGAGGAATCCTATCCATATGTAGCAGTGGATGAAATCTG
TAAGTACAGACCTGAGAATTCTGTTGCTAATGACACTGGCTTCACAGTGGTCGCACCTGGAAA
GGAGAAGGCCCTGATGAAAGCAGTCGCAACTGTGGGGCCCATCTCCGTTGCTATGGATGCAGG
CCATTCGTCCTTCCAGTTCTACAAATCAGGCATTTATTTTGAACCAGACTGCAGCAGCAAAAA
CCTGGATCATGGTGTTCTGGTGGTTGGCTACGGCTTTGAAGGAGCAAATTCGAATAACAGCAA
GTATTGGCTCGTCAAAAACAGCTGGGGTCCAGAATGGGGCTCGAATGGCTATGTAAAAATAGC
CAAAGACAAGAACAACCACTGTGGAATCGCCACAGCAGCCAGCTACCCCAATGTG**TGA**GCTGA
TGGATGGTGAGGAGGAAGGACTTAAGGACAGCATGTCTGGGGAAATTTTATCTTGAAACTGAC
CAAACGCTTATTGTGTAAGATAAACCAGTTGAATCATGGAGGATCCAAGTTGAGATTTTAATT
CTGTGACATTTTTACAAGGGTAAAATGTTACCACTACTTTAATTATTGTTATACACAGCTTTA
TGATATCAAAGACTCATTGCTTAATTCTAAGACTTTTGAATTTTCATTTTTTAAAAAGATGTA
CAAAACAGTTTGAAATAAATTTTAATTCGTATATA
```

FIGURE 12

MNLSLVLAAFCLGIASAVPKFDQNLDTKWYQWKATHRRLYGANEEGWRRAVWEKNMKMIELHN

GEYSQGKHGFTMAMNAFGDMTNEEFRQMMGCFRNQKFRKGKVFREPLFLDLPKSVDWRKKGYV

TPVKNQKQCGSCWAFSATGALEGQMFRKTGKLVSLSEQNLVDCSRPQGNQGCNGGFMARAFQY

VKENGGLDSEESYPYVAVDEICKYRPENSVANDTGFTVVAPGKEKALMKAVATVGPISVAMDA

GHSSFQFYKSGIYFEPDCSSKNLDHGVLVVGYGFEGANSNNSKYWLVKNSWGPEWGSNGYVKI

AKDKNNHCGIATAASYPNV

Important features:

Signal sequence

amino acids 1-17

N-glycosylation sites.

amino acids 2-6, 221-225, 292-296

N-myristoylation sites.

amino acids 13-19, 93-99, 136-142, 145-151, 174-180, 177-183, 180-186, 194-200, 288-294, 324-330

Eukaryotic thiol (cysteine) proteases cysteine active site.

amino acids 132-144

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 275-286

FIGURE 13

GGCGGCGTCATGTGATCCGCTTCCCTGCTCCTTTAAGCGTCCACAGGCGGCGGAGCGGCCACA
ATCACAGCTCCGGGCATTGGGGGAACCCGAGCCGGCTGCGCCGGGGGAATCCGTGCGGGCGCC
TTCCGTCCCGGTCCCATCCTCGCCGCGCTCCAGCACCTCTGAAGTTTTGCAGCGCCCAGAAAG
GAGGCGAGGAAGGAGGGAGTGTGTGAGAGGAGGGAGCAAAAAGCTCACCCTAAAACATTTATT
TCAAGGAGAAAGAAAAAGGGGGGGCGCAAAA**ATG**GCTGGGGCAATTATAGAAACATGAGCA
CCAAGAAGCTGTGCATTGTTGGTGGGATTCTGCTCGTGTTCCAAATCATCGCCTTTCTGGTGG
GAGGCTTGATTGCTCCAGGGCCCACAACGGCAGTGTCCTACATGTCGGTGAAATGTGTGGATG
CCCGTAAGAACCATCACAAGACAAATGGTTCGTGCCTTGGGGACCCAATCATTGTGACAAGA
TCCGAGACATTGAAGAGGCAATTCCAAGGGAAATTGAAGCCAATGACATCGTGTTTTCTGTTC
ACATTCCCCTCCCCCACATGGAGATGAGTCCTTGGTTCCAATTCATGCTGTTTATCCTGCAGC
TGGACATTGCCTTCAAGCTAAACAACCAAATCAGAGAAATGCAGAAGTCTCCATGGACGTTT
CCCTGGCTTACCGTGATGACGCATTTGCTGAGTGGACTGAAATGGCCCATGAAAGAGTACCAC
GGAAACTCAAATGCACCTTCACATCTCCCAAGACTCCAGAGCATGAGGGCCGTTACTATGAAT
GTGATGTCCTTCCTTTCATGGAAATTGGGTCTGTGGCCCATAAGTTTTACCTTTTAAACATCC
GGCTGCCTGTGAATGAGAAGAAGAAATCAATGTGGGAATTGGGGAGATAAAGGATATCCGGT
TGGTGGGGATCCACCAAAATGGAGGCTTCACCAAGGTGTGGTTTGCCATGAAGACCTTCCTTA
CGCCCAGCATCTTCATCATTATGGTGTGGTATTGGAGGAGGATCACCATGATGTCCCGACCCC
CAGTGCTTCTGGAAAAAGTCATCTTTGCCCTTGGGATTTCCATGACCTTTATCAATATCCCAG
TGGAATGGTTTTCCATCGGGTTTGACTGGACCTGGATGCTGCTGTTTGGTGACATCCGACAGG
GCATCTTCTATGCGATGCTTCTGTCCTTCTGGATCATCTTCTGTGGCGAGCACATGATGGATC
AGCACGAGCGGAACCACATCGCAGGGTATTGGAAGCAAGTCGGACCCATTGCCGTTGGCTCCT
TCTGCCTCTTCATATTTGACATGTGTGAGAGAGGGGTACAACTCACGAATCCCTTCTACAGTA
TCTGGACTACAGACATTGGAACAGAGCTGGCCATGGCCTTCATCATCGTGGCTGGAATCTGCC
TCTGCCTCTACTTCCTGTTTCTATGCTTCATGGTATTTCAGGTGTTTCGGAACATCAGTGGGA
AGCAGTCCAGCCTGCCAGCTATGAGCAAAGTCCGGCGGCTACACTATGAGGGGCTAATTTTTA
GGTTCAAGTTCCTCATGCTTATCACCTTGGCCTGCGCTGCCATGACTGTCATCTTCTTCATCG
TTAGTCAGGTAACGGAAGGCCATTGGAAATGGGGCGGCGTCACAGTCCAAGTGAACAGTGCCT
TTTTCACAGGCATCTATGGGATGTGGAATCTGTATGTCTTTGCTCTGATGTTCTTGTATGCAC
CATCCCATAAAAACTATGGAGAAGACCAGTCCAATGGCGATCTGGGTGTCCATAGTGGGGAAG
AACTCCAGCTCACCACCACTATCACCCATGTGGACGGACCCACTGAGATCTACAAGTTGACCC
GCAAGGAGGCCCAGGAG**TAG**GAGGCTGCAGCGCCCGGCTGGGACGGTCTCTCCATACCCCAGC
CCCTCTAACTAGAGTGGGGAGCATGCCAGAGAGAGCTCAATGTACAAATGAATGCCTCATGGC
TCTTAGCTGTGGTTTCTTGGACCAGCGGCATGGACATTTGTCAGTTTGCCTTCTGACGGTAGC
TTTTGGAGGAAGATTCCTGCAGCCACTAATGCATTGTGTATGATAACAAAAACTCTGGTATGA
CACATTTTCTGTGATCATTGTTAATTAGTGACATAGTAACATCTGTAGCAGCTGGTTAGTAAA
CCTCATGTGGGGGTGGGGTGGGGGTGTATTCCTTGGGGGATGGTTTGGGCCGAATGGGGAGTG
GAATATTTGACATTTTTCCTGTTTTAAATTCTAGGATAGATTTTAACATCCTTTGCGGTCCCA
GTCCAAGGTAGGCTGGTGTCATAGTCTTCTCACTCCTAATCCATGACCACTGTTTTTTTCCTA
TTTATATCACCAGGTAGCCTACTGAGTTAATATTTAAGTTGTCAATAGATAAGTGTCCCTGTT
TTGTGGCATAATATAACTGAATTTCATGAGAAGATTTATTCCACCAGGGGTATTTCAGCTTTG
AAACCAAATCTGTGTATCTAATACTAACCAATCTGTTGGATGTGGATTTTAAAAAATGTTTGC
TAAACTACCCAAGTAAGATTTACTGTATTAAATGGCCTTCGGGTCTGAAAGCTTTTTTAACC
TCTTGCTTAAAATGCGTTTTATTTTGATAAGATACTTCAAATAGCCTCCAAAAGTGTAGATCC
AATCACTTAAATAAACCTGTATGTATATGCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 14

MAGAIIENMSTKKLCIVGGILLVFQIIAFLVGGLIAPGPTTAVSYMSVKCVDARKNHHKTKWF
VPWGPNHCDKIRDIEEAIPREIEANDIVFSVHIPLPHMEMSPWFQFMLFILQLDIAFKLNNQI
RENAEVSMDVSLAYRDDAFAEWTEMAHERVPRKLKCTFTSPKTPEHEGRYYECDVLPFMEIGS
VAHKFYLLNIRLPVNEKKKINVGIGEIKDIRLVGIHQNGGFTKVWFAMKTFLTPSIFIIMVWY
WRRITMMSRPPVLLEKVIFALGISMTFINIPVEWFSIGFDWTWMLLFGDIRQGIFYAMLLSFW
IIFCGEHMMDQHERNHIAGYWKQVGPIAVGSFCLFIFDMCERGVQLTNPFYSIWTTDIGTELA
MAFIIVAGICLCLYFLFLCFMVFQVFRNISGKQSSLPAMSKVRRLHYEGLIFRFKFLMLITLA
CAAMTVIFFIVSQVTEGHWKWGGVTVQVNSAFFTGIYGMWNLYVFALMFLYAPSHKNYGEDQS
NGDLGVHSGEELQLTTTITHVDGPTEIYKLTRKEAQE

Important features of the protein:
Signal peptide:
amino acids 1-42

Transmembrane domains:
amino acids 239-253, 269-284, 302-318, 338-352, 377-399, 434-452,
471-488

N-glycosylation sites.
amino acids 8-12, 406-410

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 254-258

N-myristoylation sites.
amino acids 223-229, 274-280, 305-311, 358-364, 374-380, 386-392,
509-515

FIGURE 15

GTGAGGGGAACAGCTGATCCGTCTGTTGGGAGGACAGATATCTCAAGGCCAGG**ATG**GAAGAAT
CACCACTAAGCCGGGCACCATCCCGTGGTGGAGTCAACTTTCTCAATGTAGCCCGGACCTACA
TCCCCAACACCAAGGTGGAATGTCACTACACCCTTCCCCCAGGCACCATGCCCAGTGCCAGTG
ACTGGATTGGCATCTTCAAGGTGGAGGCTGCCTGTGTTCGGGATTACCACACATTTGTGTGGT
CTTCCGTGCCTGAAAGTACAACTGATGGTTCCCCCATTCACACCAGTGTCCAGTTCCAAGCCA
GCTACCTGCCCAAACCAGGAGCTCAGCTCTACCAGTTCCGATATGTGAACCGCCAGGGCCAGG
TGTGTGGGCAGAGCCCCCCTTTCCAGTTCCGAGAGCCAAGGCCCATGGATGAACTGGTGACCC
TGGAGGAGGCTGATGGGGGCTCTGACATCCTGCTGGTTGTCCCCAAGGCAACTGTGTTACAGA
ACCAGCTCGATGAGAGCCAGCAAGAACGGAATGACCTGATGCAGCTGAAGCTACAGCTGGAGG
GACAGGTGACAGAGCTGAGGAGCCGAGTGCAGGAGCTCGAGAGGGCTCTGGCAACTGCCAGGC
AGGAGCACACGGAGCTGATGGAACAGTACAAGGGGATTTCCCGGTCCCATGGGGAGATCACAG
AAGAGAGGGACATCCTGAGCCGGCAACAGGGAGACCATGTGGCACGCATCCTGGAGCTAGAGG
ATGACATCCAGACCATCAGTGAGAAAGTGCTGACGAAGGAAGTGGAGCTGGACAGGCTTAGAG
ACACAGTGAAGGCCCTGACTCGGGAACAAGAGAAGCTCCTTGGGCAACTGAAAGAAGTACAAG
CAGACAAGGAGCAAAGTGAGGCTGAGCTCCAAGTGGCACAACAGGAGAACCATCACTTAAATT
TGGACCTGAAGGAGGCGAAGAGCTGGCAAGAGGAGCAGAGTGCTCAGGCTCAGCGACTGAAAG
ACAAGGTGGCCCAGATGAAGGACACCCTAGGCCAGGCCCAGCAGCGGGTGGCCGAGCTGGAGC
CCTTGAAGGAGCAGCTTCGAGGGGCCCAGGAGCTTGCAGCCTCAAGCCAGCAGAAAGCCACCC
TTCTTGGGGAGGAGTTGGCCAGTGCAGCAGCAGCCAGGGACCGCACCATAGCCGAACTACACC
GCAGCCGCCTGGAAGTGGCTGAAGTTAACGGCAGGCTGGCTGAGCTCGGTTTGCACTTGAAGG
AAGAAAAATGCCAATGGAGCAAGGAGCGGGCAGGGCTGCTGCAGAGTGTGGAGGCAGAGAAGG
ACAAGATCCTGAAGCTGAGTGCAGAGATACTTCGATTGGAGAAGGCAGTTCAGGAGGAGAGGA
CCCAAAACCAAGTGTTCAAGACTGAGCTGGCCCGGGAGAAGGATTCTAGCCTGGTACAGTTGT
CAGAAAGTAAGCGGGAGCTGACAGAGCTGCGGTCAGCCCTGCGTGTGCTCCAGAAGGAAAAGG
AGCAGTTACAGGAGGAGAAACAGGAATTGCTAGAGTACATGAGAAAGCTAGAGGCCCGCCTGG
AGAAGGTGGCAGATGAGAAGTGGAATGAGGATGCCACCACAGAGGATGAGGAGGCCGCTGTGG
GGCTGAGCTGCCCGGCAGCTCTGACAGACTCAGAGGACGAGTCCCCAGAAGACATGAGGCTCC
CACCCTATGGCCTTTGTGAGCGTGGAGACCCAGGCTCCTCTCCTGCTGGGCCTCGAGAGGCTT
CTCCCCTTGTTGTCATCAGCCAGCCGGCTCCCATTTCTCCTCACCTCTCTGGGCCAGCTGAGG
ACAGTAGCTCTGACTCGGAGGCTGAAGATGAGAAGTCAGTCCTGATGGCAGCTGTGCAGAGTG
GGGGTGAGGAGGCCAACTTACTGCTTCCTGAACTGGGCAGTGCCTTCTATGACATGGCCAGTG
GCTTTACAGTGGGTACCCTGTCAGAAACCAGCACTGGGGGCCCTGCCACCCCCACATGGAAGG
AGTGTCCTATCTGTAAGGAGCGCTTTCCTGCTGAGAGTGACAAGGATGCCCTGGAGGACCACA
TGGATGGACACTTCTTTTTCAGCACCCAGGACCCCTTCACCTTTGAG**TGA**TCTTACTCCCTCG
TACATGCACAAATACACACTCATGCACACACACACTCACACACATGCATACACTTAGGTTTCA
TGCCCATTTTCTATCACACTGGGCTCCATGATATTCTGTTCCCTAAGAACTGCTTCTGTGTGC
CCTGTTTTCATCCCAAGATTTCTCACTTCATCCTCTCCTACCTGGCTCTTTTGTCCCAGGGAG
GGGTCCTGTTCGGAAGCAGTGGCTGAATTTATCCCCTGAAAGTGGTTTTGGAGGAACCGGGAT
GGAGGAGGCCTTCCCCTGTGGGAATAGAATCGTCCACTCCTAGCCCTGGTTGCTTCTGATACA
CAGCCACTGCACACACACTCACACTCACACTCCCTTGTCTGATGCCCCAAAGCCAATTCCT
GGGGCACCCTACCCTCTCTTATTTGGAGTTTCCGTTGGTTTACCTGAGTTTTCTCTGGGGTCT
GCACAGAGGCAGCAGCATGGACATCATGGCCTCTCAGGTCCCTTTTGGTTCTCAGTTTCATTG
GTTCCTCTTTCTGTTCCCCCATTGACTTCTGTGCCCCACCCTAGCCTTTTCCATAACCTTAGG
TATTCAGTTTGGAGGGGTTTTTTGTATTTTTGAGGATTCCTGTATTCTGTATCCTCTCCTCGC
ATCTCCTCACATGGAAAGAAATAATGTATTTGTGCCTTCTGTGAGGAATGGGGGGAACAAGTG
GTCCCAGGTATCCCCATTTCCAAGGCCCCCCTCCCTCTCCAGGTCCCCCCACAGCAATAAAAG
CTTCCCCCTGATATCCATCCCTTTGTAGTTTGAACAAATATATTTATATGATATGTAA

FIGURE 16

MEESPLSRAPSRGGVNFLNVARTYIPNTKVECHYTLPPGTMPSASDWIGIFKVEAACVRDYHT

FVWSSVPESTTDGSPIHTSVQFQASYLPKPGAQLYQFRYVNRQGQVCGQSPPFQFREPRPMDE

LVTLEEADGGSDILLVVPKATVLQNQLDESQQERNDLMQLKLQLEGQVTELRSRVQELERALA

TARQEHTELMEQYKGISRSHGEITEERDILSRQQGDHVARILELEDDIQTISEKVLTKEVELD

RLRDTVKALTREQEKLLGQLKEVQADKEQSEAELQVAQQENHHLNLDLKEAKSWQEEQSAQAQ

RLKDKVAQMKDTLGQAQQRVAELEPLKEQLRGAQELAASSQQKATLLGEELASAAAARDRTIA

ELHRSRLEVAEVNGRLAELGLHLKEEKCQWSKERAGLLQSVEAEKDKILKLSAEILRLEKAVQ

EERTQNQVFKTELAREKDSSLVQLSESKRELTELRSALRVLQKEKEQLQEEKQELLEYMRKLE

ARLEKVADEKWNEDATTEDEEAAVGLSCPAALTDSEDESPEDMRLPPYGLCERGDPGSSPAGP

REASPLVVISQPAPISPHLSGPAEDSSSDSEAEDEKSVLMAAVQSGGEEANLLLPELGSAFYD

MASGFTVGTLSETSTGGPATPTWKECPICKERFPAESDKDALEDHMDGHFFFSTQDPFTFE

Important features:

Casein kinase II phosphorylation sites:

amino acids 28-31, 43-46, 68-71, 72-75, 129-132, 156-159, 208-211, 239-242, 282-285, 305-308, 376-379, 383-383, 468-471, 520-523, 521-524, 537-540, 539-542, 543-546, 593-596, 595-598, 597-600, 612-615, 639-642, 652-655, 667-670, 683-686

N-myristoylation sites:

amino acids 39-44, 107-112, 204-209, 414-419, 561-566, 613-618

Cell attachment sequence:

amino acids 557-559

Leucine zipper pattern sequence:

amino acids 163-184, 475-496, 482-503

FIGURE 17

GCAAGTTGGGAATTTTAGACTGTCACTGCACATGGACCTCTGGGAAGACGTCTGGCGAGAGCT

AGGCCCACTGGCCCTACAGACGGATCTTGCTGGCTCACCTGTCCCTGTGGAGGTTCCCCTGGG

AAGGCAAG**ATG**CCCAACAACAGCACTGCTCTGTCATTGGCCAATGTTACCTACATCACCATGG

AAATTTTCATTGGACTCTGCGCCATAGTGGGCAACGTGCTGGTCATCTGCGTGGTCAAGCTGA

ACCCCAGCCTGCAGACCACCACCTTCTATTTCATTGTCTCTCTAGCCCTGGCTGACATTGCTG

TTGGGGTGCTGGTCATGCCTTTGGCCATTGTTGTCAGCCTGGGCATCACAATCCACTTCTACA

GCTGCCTTTTTATGACTTGCCTACTGCTTATCTTTACCCACGCCTCCATCATGTCCTTGCTGG

CCATCGCTGTGGACCGATACTTGCGGGTCAAGCTTACCGTCAGATTCAGAATTCCTGGGCTCC

CTGGGTGCATTCTATCATTCCAGTTGAAAGTTTGCTTCCTTCCAGTCATGTGGCTCTTCATTC

TACTCCTTGGCTCTCATTTCAGATGCCATGGTCATGGATGAAAAGGTCAAGAGAAGCTTTG

TGCTGGACACGGCTTCTGCCATCTGCAACTACAATGCCCACTACAAGAATCACCCCAAATACT

GGTGCCGAGGCTATTTCCGTGACTACTGCAACATCATCGCCTTCTCCCCTAACAGCACCAATC

ATGTGGCCCTGAGGGACACAGGGAACCAGCTCATTGTCACTATGTCCTGCCTGACCAAAGAGG

ACACGGGCTGGTACTGGTGTGGCATCCAGCGGGACTTTGCCAGGGATGACATGGATTTTACAG

AGCTGATTGTAACTGACGACAAAGGAACCCTGGCCAATGACTTTTGGTCTGGGAAAGACCTAT

CAGGCAACAAAACCAGAAGCTGCAAGGCTCCCAAAGTTGTCCGCAAGGCTGACCGCTCCAGGA

CGTCCATTCTCATCATTTGCATACTGATCACGGGTTTGGGAATCATCTCTGTAATCAGTCATT

TGACCAAAAGGAGGAGAAGTCAAAGGAATAGAAGGGTAGGCAACACTTTGAAGCCCTTCTCGC

GTGTCCTGACTCCAAAGGAAATGGCTCCTACTGAACAGATG**TGA**CTGAAGATTTTTTAATTT

AGTTCATAAAGTGATGCTACAACAGAATAATCACCATGACAACTGGCCCACACCTCAGAGACT

GATTCTGATCTCCCAGGAATTCTGAAGGACCCTCTATCCTTGACAACAATCATTTGCAGCCAG

GTAGCAACGGCGGTAGTCAGAGGAGCTATGATAGACCACACCCAAGCAAGGCTGCCCTCAAAT

AACATCTCAAGATCTTAGTTCTTATGCATTCCATCAGTCAGAAGTGAAGAAGAGGTGGAGAAT

CTGGATTGGGGACCAGGAAATCACTTGTATTTTGTTAGCCAATAAATTCCTAGCCAGTGTTGA

ATGAAAAAAAAAAAAA

FIGURE 18

MPNNSTALSLANVTYITMEIFIGLCAIVGNVLVICVVKLNPSLQTTTFYFIVSLALADIAVGV
LVMPLAIVVSLGITIHFYSCLFMTCLLLIFTHASIMSLLAIAVDRYLRVKLTVRFRIPGLPGC
ILSFQLKVCFLPVMWLFILLSLALISDAMVMDEKVKRSFVLDTASAICNYNAHYKNHPKYWCR
GYFRDYCNIIAFSPNSTNHVALRDTGNQLIVTMSCLTKEDTGWYWCGIQRDFARDDMDFTELI
VTDDKGTLANDFWSGKDLSGNKTRSCKAPKVVRKADRSRTSILIICILITGLGIISVISHLTK
RRRSQRNRRVGNTLKPFSRVLTPKEMAPTEQM

Important features of the protein:

Transmembrane domains:

amino acids 16-35, 62-80, 89-101, 134-152, 292-311

N-glycosylation sites.

amino acids 3-7, 4-8, 12-16, 204-208, 273-277

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 316-320

N-myristoylation sites.

amino acids 122-128, 125-131, 258-264

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 214-225

G-protein coupled receptors proteins.

amino acids 29-59, 76-116

FIGURE 19

```
CTCGGGCGCGCACAGGCAGCTCGGTTTGCCCTGCGATTGAGCTGCGGGTCGCGGCCGGCGCCGGCCTCTCCAATG
GCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTCGGCAAAGGCAGTCGAGTGTTTGCAGACCGGGGCGAGTC
CTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCGGCCGGGGCTGTCGCACT
CCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTCACGTCG
TTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCCCCCTG
GTTTGTGTGTTACGCACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCCTGGGCG
AATCCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTCGAATCTGCGAGTGAAGAGGG
ACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGATCAGCAAAA
AAAGAAG**ATG**GGCCCCCCGAGCCTCGTGCTGTGCTTGCTGTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTC
GGCCTTCCTGTCGCACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCAACATCATCCT
GGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGCATCATGGAGCA
GGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCTCACGCTCCTCCATCCTCACTGG
CAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGCAGGCACAGCACGA
GAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATA
CAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTCGGACTCCTTAAAAACTCCCGCTTTTATAACTACAC
GCTGTGTCGGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCACAGACCTCATCACCAA
TGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCCTCATGGTCATCAGCCATGC
AGCCCCCACGGCCCTGAGGATTCAGCCCCCACAATATTCACGCCTCTTCCCAAACGCATCTCAGCACATCACGCC
GAGCTACAACTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACAT
GGAATTCACCAACATGCTCCAGCGGGAAGCGCTTGCAGACCCTCATGTCGGTGGACGACTCCATGGAGACGATTTA
CAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGCCGACCACGGTTACCACATCGG
CCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTGAGGGGCCCC
CAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCCTCAACATTGACCTGGCCCCCACCATCCTGGACATTGC
AGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCTGGACACGGAGCGGCCGGTGAATCG
GTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGGAGAGAGGCAAGCTGCTACACAAGAG
AGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTACCAGCGTGTGAAGGACCTGTGTCAGCG
TGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTGTGTGGAGGACGCCACGGGGAAGCTGAA
GCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCCTCTCCAACCTCGTGCCCAAGTACTACGG
GCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGGACGCCGGAAAAAACTCTT
CAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGACGGCAGGGT
GTACCACGTAGGCCTGGGTGATGCCGCCCAGCCCCGAAACCTCACCAAGCGGCACTGGCCAGGGGCCCCTGAGGA
CCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGCCTTCCCGACTACTCAGCCGCCAACCCCATTAA
AGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGC
CTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAACCCTGCAGAACAAAATTAAGAACCTGAGGGAAGT
CCGAGGTCACCTGAAGAAAAAGCGGCCAGAAGAATGTGACTGTCACAAAATCAGCTACCACACCCAGCACAAAGG
CCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAGGAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTT
GCGGGAGCAGAAGCGCAAGAAGAAACTCCGCAAGCTGCTCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCC
AGGCCTCACGTGCTTCACCCACGACAACCAGCACTGGCAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGC
CTGCACCAGCGCCAACAATAACACGTACTGGTGCATGAGGACCATCAATGAGACTCACAATTTCCTCTTCTGTGA
ATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACT
GGACAGGGATGTCCTCAACCAGCTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAA
CCCCCGGACTCGAAACATGGACCTGGATGGAGGAAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCC
AGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGGGAAGGCTGGGAAGGT**TAA**GAAACAACAGAGGT
GGACCTCCAAAAACATAGAGGCATCACCTGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACC
TGTGCTATTGGCCAGGAGGCCTGAGAAAGCAAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCA
GCAGGAGCAGAGATAACTTCAGGAAGTCCATTTTTGCCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCAC
AAAATGCATTTTTTCGTATCAAAAAGTCACCACTAACCCTCCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCG
AGCGAGAGAGATTTCCTTGGAAATTTCTCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCA
GTCCTGTTCTAAATCCTCTTATTCTTTTGGTTTGTCACAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGG
AGAGGCTGAAAACAGTGCAGAGACGTTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTAT
AAACCCTGGTTGCCTCTGAAGAAACTGCCTTCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACT
TTTAGGGGAACCTAATAAGAAATCCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAA
GAAAAA
```

FIGURE 20

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQV

MNKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHESRT

FAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSDYSK

DYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITPSYNY

APNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNTYIVYT

ADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDIAGLDIP

ADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEENFLPKYQR

VKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKYYGQGSEAC

TCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVAIEVDGRVYHVGLGDAAQPRNLTKRHW

PGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLH

IDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSLHPFRKGLQEKD

KVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNNT

YWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGY

KQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

Important features:

Signal peptide:

amino acids 1-17

Sulfatases signature 1.

amino acids 86-99

Homologous region to sulfatase:

amino acids 87-106, 133-146, 216-229, 291-320, 365-375

N-glycosylation sites.

amino acids 65-69, 112-116, 132-136, 149-153, 171-175, 198-202, 241-245, 561-565, 608-612, 717-721, 754-758, 764-768

FIGURE 21

GGGCGCGCGAGAGCTGCTAGGGCGGTTTCTCTGCCTCGGGCCTGTTGGGCAGGGCCGGCT

AAGGTGCGCGTGCTCGCTGGTTCTAACCCTTCTGTTGGGCGTTTCTGCTGAGAGGCGGGA

GGCGCTGAGAGTCTGTGCGGAGGTCCGTGGACAGACTGCTTTGCTCGTTGTTGCTCTTCG

GAGGCGGCGATCCCCGAAGGCGAGCTGAAATACGGCTGCAGGCTACAATTTGCAGCCGAC

GATTATGGAAGACGGAAGCGGGAGAGGTGGCCCACCCTC**ATG**GAGCGCTTGTGCTCGGAT

GGCTTCGCATTTCCCCAATACCCCATTAAACCGTATCATCTGAAGAGGATCCACAGAGCT

GTCTTACATGGTAATCTAGAGAAACTGAAGTACCTTCTGCTCACGTATTATGACGCCAAT

AAGAGAGACAGGAAGGAAAGGACCGCCCTACATTTGGCCTGTGCCACTGGCCAACCGGAA

ATGGTACATCTCCTGGTGTCCAGAAGATGTGAGCTTAACCTCTGCGACCGTGAAGACAGG

ACACCTCTGATCAAGGCTGTACAACTGAGGCAGGAGGCTTGTGCAACTCTTCTGCTGCAA

AATGGCGCCAATCCAAATATTACGGATTTCTTTGGAAGGACTGCTCTGCACTACGCTGTG

TATAATGAAGATACATCCATGATAGAAAACTTCTTTCACATGGTACAAATATTGAAGAA

TGCAGCAAGGTA**TAG**GTCAACCAATGTTATTTTCAAACTATCTGAAATGAATTTATTTTA

ACATTGACACATGTAAGGGTCAATTTTTCATATTTGGAAGCTCAAACATTCCTTGAATGA

AAATATTTTGAAATGCCTTAACTGTCTAAGATTTTACTTTAAATATTGGAACTTTTAAAG

AAGCATTATAGGGAACAGCCTTTTTTTCATGCACTTATGGTAAATAACTATAAAAACAAAT

GAATTACAATAAATTTATAATTCATGACAACTGAATTTGGGAAAGGTAATAGTTAAGTGT

TTTTCCACTAAATTACTTTTT

FIGURE 22

MERLCSDGFAFPQYPIKPYHLKRIHRAVLHGNLEKLKYLLLTYYDANKRDRKERTALHLACAT
GQPEMVHLLVSRRCELNLCDREDRTPLIKAVQLRQEACATLLLQNGANPNITDFFGRTALHYA
VYNEDTSMIEKLLSHGTNIEECSKV

Important features of the protein:

N-glycosylation site.

amino acids 113-117

N-myristoylation site.

amino acids 109-115

Microbodies C-terminal targeting signal.

amino acids 149-153

FIGURE 23

GAGGCAGAAAGGCAGAAAGGAGAAAATTCAGGATAACTCTCCTGAGGGGTGAGCCAAGCCCTG

CCATGTAGTGCACGCAGGACATCAACAAACACAGATAACAGGAAATGATCCATTCCCTGTGGT

CACTTATTCTAAAGGCCCCAACCTTCAAAGTTCAAGTAGTGAT**ATG**GATGACTCCACAGAAAG

GGAGCAGTCACGCCTTACTTCTTGCCTTAAGAAAAGAGAAGAAATGAAACTGAAGGAGTGTGT

TTCCATCCTCCCACGGAAGGAAAGCCCCTCTGTCCGATCCTCCAAAGACGGAAAGCTGCTGGC

TGCAACCTTGCTGCTGGCACTGCTGTCTTGCTGCCTCACGGTGGTGTCTTTCTACCAGGTGGC

CGCCCTGCAAGGGGACCTGGCCAGCCTCCGGGCAGAGCTGCAGGGCCACCACGCGGAGAAGCT

GCCAGCAGGAGCAGGAGCCCCCAAGGCCGGCCTGGAGGAAGCTCCAGCTGTCACCGCGGGACT

GAAAATCTTTGAACCACCAGCTCCAGGAGAAGGCAACTCCAGTCAGAACAGCAGAAATAAGCG

TGCCGTTCAGGGTCCAGAAGAAACAGTCACTCAAGACTGCTTGCAACTGATTGCAGACAGTGA

AACACCAACTATACAAAAAGGATCTTACACATTTGTTCCATGGCTTCTCAGCTTTAAAAGGGG

AAGTGCCCTAGAAGAAAAGAGAATAAAATATTGGTCAAAGAAACTGGTTACTTTTTTATATA

TGGTCAGGTTTTATATACTGATAAGACCTACGCCATGGGACATCTAATTCAGAGGAAGAAGGT

CCATGTCTTTGGGGATGAATTGAGTCTGGTGACTTTGTTTCGATGTATTCAAAATATGCCTGA

AACACTACCCAATAATTCCTGCTATTCAGCTGGCATTGCAAAACTGGAAGAAGGAGATGAACT

CCAACTTGCAATACCAAGAGAAAATGCACAAATATCACTGGATGGAGATGTCACATTTTTTGG

TGCATTGAAACTGCTG**TGA**CCTACTTACACCATGTCTGTAGCTATTTTCCTCCCTTTCTCTGT

ACCTCTAAGAAGAAAGAATCTAACTGAAAATACCAAAAAAAAAAAAAAA

FIGURE 24

MDDSTEREQSRLTSCLKKREEMKLKECVSILPRKESPSVRSSKDGKLLAATLLLALLSCCLTV
VSFYQVAALQGDLASLRAELQGHHAEKLPAGAGAPKAGLEEAPAVTAGLKIFEPPAPGEGNSS
QNSRNKRAVQGPEETVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKRGSALEEKENKILVKE
TGYFFIYGQVLYTDKTYAMGHLIQRKKVHVFGDELSLVTLFRCIQNMPETLPNNSCYSAGIAK
LEEGDELQLAIPRENAQISLDGDVTFFGALKLL

Transmembrane domain:
amino acids 47-72

N-glycosylation site.
amino acids 124-127, 242-245

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36, 173-176

N-myristoylation site.
amino acids 96-101

TNF family proteins.
amino acids 172-206

FIGURE 25

CTGCTTGGATACCTCCAGTCCCCAAACTGTGTTCCAGGAGTTTTCTTGGCCGAAGCTGCCCGA
TGTTTGAGCCTTTTCTTCCCAGAGAAGAAGATGGACTGAAAGCTGCCAGTTGGGGACTTTTTG
TGATCACGGCGTTGCAGCGTTTTAAAGGAGGTGATGGGGCTTGCGCTGGCTTGTCTTCCCACC
CAAGTGAAGAGTTGATGTTCACTGGTTATGCTTAGACAATGTGCAGTTTGTGTTAATTTAAAA
TTTTGGGTGGGATAGGGGCATAGGCTTGTGAAGGGCAGTCCGGATCCGGAGGAACTCGTCTTT
GTCCCTGGTAGGAGAGACACCCCCAGTCTATCCTCGATGCCGTCAGCCTTGGCCATCTTCACT
TGCCGCCCGAACTCGCACCCGTTTCAGGAGCGTCATGTCTACCTGGACGAGCCCATCAAAATC
GGCCGCTCAGTGGCCCGCTGTCGACCAGCGCAGAATAATGCCACTTTTGATTGCAAAGTGCTA
TCAAGGAACCACGCTCTCGTCTGGTTTGATCACAAGACGGGCAAGTTTTATCTTCAAGACACT
AAAAGTAGTAATGGTACTTTTATAAATAGCCAGAGATTGAGTCGAGGCTCTGAAGAAAGTCCA
CCATGTGAAATTCTTTCCGGTGACATTATCCAGTTTGGAGTAGACGTGACAGAGAATACACGG
AAAGTTACCCATGGGTGTATTGTTTCCACAATAAAACTTTTTCTACCAGATGG**ATG**GAAGCC
CGGCTCCGCTCAGATGTCATCCATGCACCATTACCAAGTCCTGTTGACAAAGTTGCTGCTAAC
ACTCCAAGTATGTACTCTCAGGAACTATTCCAGCTTTCTCAGTATCTACAGGAGGCCTTACAT
CGGGAACAAATGTTGGAACAGAAGTTAGCCACGCTTCAGCGGCTACTAGCCATCACCCAAGAG
GCTTCAGATACCAGTTGGCAGGCTTTAATAGATGAAGATAGACTCTTATCACGGTTAGAAGTT
ATGGGAAACCAATTACAGGCATGCTCCAAAAATCAAACAGAAGATAGTTTACGAAAGGAACTT
ATAGCATTACAAGAGGATAAACATAACTATGAGACAACAGCCAAAGAGTCCCTGAGGCGGGTT
CTTCAGGAGAAAATTGAAGTGGTTAGAAACTTTCAGAAGTTGAGCGAAGTCTGAGTAATACT
GAAGATGAATGTACCCATCTGAAAGAAATGAATGAAAGGACTCAGGAAGAATTAAGAGAATTA
GCCAACAAATATAATGGAGCAGTTAATGAGATTAAAGATTTATCTGATAAATTAAAGGTAGCA
GAGGGAAAACAAGAGGAAATCCAACAGAAGGGACAGGCTGAGAAAAAGAATTACAACATAAA
ATAGATGAAATGGAAGAAAAAGAACAGGAGCTCCAGGCAAAAATAGAAGCTTTGCAAGCTGAT
AATGATTTCACCAATGAAAGGCTAACAGCTTTACAAGTACGGTTAGAACATCTTCAGGAGAAA
ACTCTTAAAGAATGCAGCAGCTTGGCTGATCGTCGAAGGGCATCTAACCAAAGCGGTAGAAGA
AACAAAGCTTTCAAAAGGTTTGTTTTCTGTTTTTCTATGTTTTTTGACAGTTCTTTTGGA**TAA**
TGAAGGTTAGTGTATATTTTCAAGGTTATAGTATTTTAACCATCAGTTTACTTCTTATAGCTC
ACAAAATAGCAAGCCAGTAACAGTATCAGATAATATATAAAATAATCAGACTTCTGTTTTAAG
AAGGGTATCGTAACTGGAATGTGTCTTTTTAAGTGGATGTATATTTATGGTTTTTTGAATGTT
AGTACTTGATATAGGTTTCTTTAGGTATTAAAGATTTGTTGCAATCTCTGTCATTCCCAGCAT
TAATTTCAGCTTTGATCTCAAATTTTAATCAAACACAATGTAAGTCGTTTGTGATACAACTTA
AGTGAAACATGCTTGCACTTCTATTTTGGGGGTTACAGTACCTTTAAAATCTCTTATGATGTT
TAATATTTCCTTAATTTTTGGCATCTCAGTTTGATTTAAACAAAATTAATGACTTTTGTGAAT
GTAGAATCTTCTTATATTTTATGAGTAGTCCAGTAATTGCCCAAAGTAGTTTATTGTGTTAAT
TCTGTTACAGTTGTCAGAGAAGAAAGTGAGTTTTAAAGCACCATATTGTCAAGTCACTTTTA
TACATAGGGAAATTAGGCAAATAAATTTGGTGGCATGTGTTTATCATAGTAGAACTTTCATTA
GACTATACCAGTATAAAATTTAAAACTAGATTCACAGTCCTTTTGGCCAATTAAAACATTGAG
TTACAAAGTTTGAGATACTTAATTTTAGTACATTCTATTTTATTAAAGTAACTGGATTCATT
TGACTTTTTTAACCATGTAAGAGGATGGTGTTATTTCAAATATCTCGTGGTTTCCATTCTGAA
TTTTGTGCACGGCAGATGCCATATTTGGGGAAAAAATGCATAGAATATGCATCATTAATATTG
TTTTGGCAAACAGGCATTGAGTTTCAGAACAGTGAACTATTTTTAGTACATATGGCAATTTTT
TTCACCTTATTAAAGTGAGATGAGAACAGACCTTAAAATAGCTTTTACCTCACCATCCAAATA
CCTATTCAGATTAGTTGGTTGAATAGCCAGCACTTTGAAGTAGAGCCTTAGG

FIGURE 26

MEARLRSDVIHAPLPSPVDKVAANTPSMYSQELFQLSQYLQEALHREQMLEQKLATLQRLLAI
TQEASDTSWQALIDEDRLLSRLEVMGNQLQACSKNQTEDSLRKELIALQEDKHNYETTAKESL
RRVLQEKIEVVRKLSEVERSLSNTEDECTHLKEMNERTQEELRELANKYNGAVNEIKDLSDKL
KVAEGKQEEIQQKGQAEKKELQHKIDEMEEKEQELQAKIEALQADNDFTNERLTALQVRLEHL
QEKTLKECSSLADRRRASNQSGRRNKAFKRFVFCFSMFFDSSFG

Important features of the protein:

N-glycosylation sites.

amino acids 98-102, 271-275

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 138-142, 267-271

Amidation site.

amino acids 273-277

Tropomyosins proteins.

amino acids 169-217

FIGURE 27

GAACCTGGCGCCGCCGGAACTGATCGCGGCCTAGTCCCGACGCGTGTGTGCTAGTGAGCCGGA
GCCGGCGACGGCGGCAGTGGCGGCCCGGCCTGCAGGAGCCCGACGGGGTCTCTGCCATGGGGG
AGTGACGCGCCTGCACCCGCTGTTCCGCGGCAGCGGCGAGACATGAGGAGACCCCGCGACAGG
GGCAGCGGCGGCGGCTCGTGAGCCCCGGG**ATG**GAGGAGAAATACGGCGGGGACGTGCTGGCCG
GCCCCGGCGGCGGCGGCGGCCTTGGGCCGGTGGACGTACCCAGCGCTCGATTAACAAAATATA
TTGTGTTACTATGTTTCACTAAATTTTTGAAGGCTGTGGGACTTTTCGAATCATATGATCTCC
TAAAAGCTGTTCACATTGTTCAGTTCATTTTTATATTAAAACTTGGGACTGCATTTTTTATGG
TTTTGTTTCAAAAGCCATTTTCTTCTGGGAAAACTATTACCAAACACCAGTGGATCAAAATAT
TTAAACATGCAGTTGCTGGGTGTATTATTTCACTCTTGTGGTTTTTTGGCCTCACTCTTTGTG
GACCACTAAGGACTTTGCTGCTATTTGAGCACAGTGATATTGTTGTCATTTCACTACTCAGTG
TTTTGTTCACCAGTTCTGGAGGAGGACCAGCAAAGACAAGGGGAGCTGCTTTTTTCATTATTG
CTGTGATCTGTTTATTGCTTTTTGACAATGATGATCTCATGGCTAAAATGGCTGAACACCCTG
AAGGACATCATGACAGTGCTCTAACTCATATGCTTTACACAGCCATTGCCTTCTTAGGTGTGG
CAGATCACAAGGGTGGAGTATTATTGCTAGTACTGGCTTTGTGTTGTAAAGTTGGTTTTCATA
CAGCTTCCAGAAAGCTCTCTGTCGACGTTGGTGGAGCTAAACGTCTTCAAGCTTTATCTCATC
TTGTTTCTGTGCTTCTCTTGTGCCCATGGGTCATTGTTCTTTCTGTGACAACTGAGAGTAAAG
TGGAGTCTTGGTTTTCTCTCATTATGCCTTTTGCAACGGTTATCTTTTTTGTCATGATCCTGG
ATTTCTACGTGGATTCCATTTGTTCAGTCAAAATGGAAGTTTCCAAATGTGCTCGTTATGGAT
CCTTTCCCATTTTTATTAGTGCTCTCCTTTTTGGAAATTTTTGGACACATCCAATAACAGACC
AGCTTCGGGCTATGAACAAAGCAGCACACCAGGAGAGCACTGAACACGTCCTGTCTGGAGGAG
TGGTAGTGAGTGCTATATTCTTCATTTTGTCTGCCAATATCTTATCATCTCCCTCTAAGAGAG
GACAAAAAGGTACCCTTATTGGATATTCTCCTGAAGGAACACCTCTTTATAACTTCATGGGTG
ATGCTTTTCAGCATAGCTCTCAATCGATCCCTAGGTTTATTAAGGAATCACTAAAACAAATTC
TTGAGGAGAGTGACTCTAGGCAGATCTTTTACTTCTTGTGCTTGAATCTGCTTTTTACCTTTG
TGGAATTATTCTATGGCGTGCTGACCAATAGTCTGGGCCTGATCTCGGATGGATTCCACATGC
TTTTTGACTGCTCTGCTTTAGTCATGGGACTTTTTGCTGCCCTGATGAGTAGGTGGAAAGCCA
CTCGGATTTTCTCCTATGGGTACGGCCGAATAGAAATTCTGTCTGGATTTATTAATGGACTTT
TTCTAATAGTAATAGCGTTTTTTGTGTTTATGGAGTCAGTGGCTAGATTGATTGATCCTCCAG
AATTAGACACTCACATGTTAACACCAGTCTCAGTTGGAGGGCTGATAGTAAACCTTATTGGTA
TCTGTGCCTTTAGCCATGCCCATAGCCATGCCCATGGAGCTTCTCAAGGAAGCTGTCACTCAT
CTGATCACAGCCATTCACACCATATGCATGGACACAGTGACCATGGGCATGGTCACAGCCACG
GATCTGCGGGTGGAGGCATGAATGCTAACATGAGGGGTGTATTTCTACATGTTTTGGCAGATA
CACTTGGCAGCATTGGTGTGATCGTATCCACAGTTCTTATAGAGCAGTTTGGATGGTTCATCG
CTGACCCACTCTGTTCTCTTTCTACTGCTATATTAATATTTCTCAGTGTTGTTCCACTGATTA
AAGATGCCTGCCAGGTTCTACTCCTGAGATTGCCACCAGAATATGAAAAGAACTACATATTG
CTTTAGAAAGATACAGAAATTGAAGGATTAATATCATACCGAGACCCTCATTTTTGGCGTC
ATTCTGCTAGTATTGTGGCAGGAACAATTCATATACAGGTGACATCTGATGTGCTAGAACAAA
GAATAGTACAGCAGGTTACAGGAATACTTAAAGATGCTGGAGTAAACAATTAACAATTCAAG
TGGAAAAGGAGGCATACTTTCAACATATGTCTGGCCTAAGTACTGGATTTCATGATGTTCTGG
CTATGACAAAACAAATGGAATCCATGAAATACTGCAAAGATGGTACTTACATCATG**TGA**GATA
ACTCAAGAATTACCCCTGGAGAATAAACAATGAAGATTAAATGACTCAGTATTTGTAATATTG
CCAGAAGGATAAAAATTACACATTAACTGTACAGAAACAGAGTTCCCTACTACTGGATCAAGG
AATCTTTCTTGAAGGAAATTTAAATACAGAATGAAACATTAATGGTAAAAAAA

FIGURE 28

MEEKYGGDVLAGPGGGGGLGPVDVPSARLTKYIVLLCFTKFLKAVGLFESYDLLKAVHIVQFI
FILKLGTAFFMVLFQKPFSSGKTITKHQWIKIFKHAVAGCIISLLWFFGLTLCGPLRTLLLFE
HSDIVVISLLSVLFTSSGGGPAKTRGAAFFIIAVICLLLFDNDDLMAKMAEHPEGHHDSALTH
MLYTAIAFLGVADHKGGVLLLVLALCCKVGFHTASRKLSVDVGGAKRLQALSHLVSVLLLCPW
VIVLSVTTESKVESWFSLIMPFATVIFFVMILDFYVDSICSVKMEVSKCARYGSFPIFISALL
FGNFWTHPITDQLRAMNKAAHQESTEHVLSGGVVVSAIFFILSANILSSPSKRGQKGTLIGYS
PEGTPLYNFMGDAFQHSSQSIPRFIKESLKQILEESDSRQIFYFLCLNLLFTFVELFYGVLTN
SLGLISDGFHMLFDCSALVMGLFAALMSRWKATRIFSYGYGRIEILSGFINGLFLIVIAFFVF
MESVARLIDPPELDTHMLTPVSVGGLIVNLIGICAFSHAHSHAHGASQGSCHSSDHSHSHHMH
GHSDHGHGHSHGSAGGGMNANMRGVFLHVLADTLGSIGVIVSTVLIEQFGWFIADPLCSLSTA
ILIFLSVVPLIKDACQVLLLRLPPEYEKELHIALEKIQKIEGLISYRDPHFWRHSASIVAGTI
HIQVTSDVLEQRIVQQVTGILKDAGVNNLTIQVEKEAYFQHMSGLSTGFHDVLAMTKQMESMK
YCKDGTYIM

Important features of the protein:

Signal peptide:

amino acids 1-46

Transmembrane domains:

amino acids 59-77, 101-119, 150-167, 205-223, 239-258, 267-284, 305-324, 343-360, 421-440, 452-469, 486-505, 522-539, 592-612, 621-641

N-glycosylation site.

amino acids 721-725

Glycosaminoglycan attachment site.

amino acids 143-147

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 225-229

Tyrosine kinase phosphorylation sites.

amino acids 750-758, 756-764

N-myristoylation sites.

amino acids 14-20, 46-52, 102-108, 112-118, 144-150, 317-323, 347-353, 369-375, 372-378, 437-443, 462-468, 529-535, 549-555, 553-559, 579-585, 582-588, 583-589, 584-590, 605-611, 737-743

Multicopper oxidases protein:

amino acids 561-569

FIGURE 29

GGCACGAGGGCAGGATATTAGAA**ATG**GCTACTCCCCAGTCAATTTTCATCTTTGCAATCTGCA

TTTTAATGATAACAGAATTAATTCTGGCCTCAAAAAGCTACTATGATATCTTAGGTGTGCCAA

AATCGGCATCAGAGCGCCAAATCAAGAAGGCCTTTCACAAGTTGGCCATGAGTACCACCCTG

ACAAAAATAAGAGCCCGGATGCTGAAGCAAAATTCAGAGAGATTGCAGAAGCATATGAAACAC

TCTCAGATGCTAATAGACGAAAGAGTATGATACACTTGGACACAGTGCTTTTACTAGTGGTA

AAGGACAAAGAGGTAGTGGAAGTTCTTTTGAGCAGTCATTTAACTTCAATTTTGATGACTTAT

TTAAAGACTTTGGCTTTTTTGGTCAAAACCAAACACTGGATCCAAGAAGCGTTTTGAAAATC

ATTTCCAGACACGCCAGGATGGTGGTTCCAGTAGACAAAGGCATCATTTCCAAGAATTTTCTT

TTGGAGGTGGATTATTTGATGACATGTTTGAAGATATGGAGAAAATGTTTTCTTTTAGTGGTT

TTGACTCTACCAATCAGCATACAGTACAGACTGAAAATAGATTTCATGGATCTAGCAAGCACT

GCAGGACTGTCACTCAACGAAGAGGAAATATGGTTACTACATACACTGACTGTTCAGGACAG**T**

AGTTCTTATTCTATTCTCACTAAATCCAACTGGTTGACTCTTCCTCATTATCTTTGATGCTAA

ACAATTTTCTGTGAACTATTTTGACAAGTGCATGATTTCACTTTAAACAATTTGATATAGCTA

TTAAATATATTTAAGGGTTTTTTTTTTTGACAAATTCAACATTCAACGAGTAGACAAAATGCT

AATTATTTCCCTGATTAGGAAAGTTTCTTTAAAAAACACGTAATTTTGCCTAGTGCTTTTTCT

CTACCTGCCCTTGGGCTCACTAATATCACCAGTATTATTACCAAGAAATATTGAGTTTACCT

GATTAAACTTTAAAAGTTAATTGTAGATTTAAATTGTGTGAACCTAATGATTTTTGCAGTGAA

ACCTTTACTAATTCAAAGTTGCATGTTCTATGACATCTGTGACTTGCGTTGCAGAGTGTACAT

GAAACTGTATAATTGAGTCATTCAGTAAAGGAGAACAGTATCTTGGTTAATTGCTACTGAAAG

GTTGAGAAAGGAATGGTTTGATATTTACCACAGCGCTGTGCCTTTCTACAGTAGAACTGGGGT

AAAGGAAATGGTTTTATTGCCCATAGTCATTTAGGCTGGAAAAAGTTGAAAACTTAACGAAA

TATTGCCAAGAGATTGTTATGTGTTTGGTTCCAGCCTAAAAATGATTTTGTAGTGTTGAAATC

ATAGCTACTTACATAGCTTTTTCATATTTCTTTCTTAGTTGTTGGCACTCTTAGGTCTTAGTA

TGGATTTATGTGTTTGTGTGTGTGTAGTTTATCCTCTCTCTCATCTTTATCTAGAGATTGACT

GATACCTCATTCTGTTTGTAAAACCAGCCAGTAATTTCTGTGCAACCTTACTATGTGCAATAT

TTTTAAATCCTGAGAAATGTGTGCTTTTGTTTTCGGATAGACTTATTTCTTTAGTTCTGCACT

TTTCCACATTATACTCCATATGAGTATTAATCCTATGGATACATATTAAAACAAGTGTCTCAT

FIGURE 30

MATPQSIFIFAICILMITELILASKSYYDILGVPKSASERQIKKAFHKLAMKYHPDKNKSPDA
EAKFREIAEAYETLSDANRRKEYDTLGHSAFTSGKGQRGSGSSFEQSFNFNFDDLFKDFGFFG
QNQNTGSKKRFENHFQTRQDGGSSRQRHHFQEFSFGGGLFDDMFEDMEKMFSFSGFDSTNQHT
VQTENRFHGSSKHCRTVTQRRGNMVTTYTDCSGQ

Important features of the protein:

Signal peptide:

amino acids 1-23

Nt-dnaJ domain signature.

amino acids 27-59, 66-90

Glycosaminoglycan attachment site.

amino acids 96-100

N-myristoylation sites.

amino acids 32-38, 99-105, 102-108, 126-132, 211-217

FIGURE 31

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTGG

GCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACAAT

TCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCTGAG

ATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA**ATG**

CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCTACGCA

TTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCTCAGAACCTCTCTGTA

CTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGAAACAGTG

TACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCTGGATCCCC

AGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATCACGGCCACT

GTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTGGAGCATCCTG

AAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGATCACCAAAGAT

GGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTCCTTGTGGCCTAC

TGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGGGGGTATTCCAGTG

CACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGACATTCGTGAAGGCC

ATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAAGGAGAGGCCATTCCC

CTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGTGGTCGTGCCACTGTTC

GTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGGTGGTCCTCCCAGACACC

TTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGGGAGGAGGTGGATGCCTGT

GCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGATCTCA**TAG**GTTTGCGGAAGG

GCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACCATGAGGGGACAAGTTGTGTT

TCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGAGCCTGTTGTCTACAAGTCTAG

AAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACACTGACTGAGGCTTAGGGGATGTG

ACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACCCTGGGAAAAGTGACTTCATCCCT

TCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACCTACACACCTGCTAAACACACACAC

ACAGAGTCTCTCTATATATACACACGTACACATAAATACACCCAGCACTTGCAAGGCTAGA

GGGAAACTGGTGACACTCTACAGTCTGACTGATTCAGTGTTTCTGGAGAGCAGGACATAAATG

TATGATGAGAATGATCAAGGACTCTACACACTGGGTGGCTTGGAGAGCCCACTTTCCCAGAAT

AATCCTTGAGAGAAAAGGAATCATGGGAGCAATGGTGTTGAGTTCACTTCAAGCCCAATGCCG

GTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGTAGGTGACCTGGAGGAAGGTCACAGCCACA

CTGAAAATGGGATGTGCATGAACACGGAGGATCCATGAACTACTGTAAAGTGTTGACAGTGTG

TGCACACTGCAGACAGCAGGTGAAATGTATGTGTGCAATGCGACGAGAATGCAGAAGTCAGTA

ACATGTGCATGTTTGTTGTGCTCCTTTTTTTCTGTTGGTAAAGTACAGAATTCAGCAAATAAAA

AGGGCCACCCTGGCCAAAAGCGGTAAAAAAAAAAAAAAAA

FIGURE 32

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGET
VYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAWSI
LKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSGGIP
VHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILVVVPL
FVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:
Signal peptide:
amino acids 1-29

Transmembrane domain:
amino acids 230-255

N-glycosylation sites.
amino acids 40-44, 134-138

Tissue factor proteins.
amino acids 92-120

Integrins alpha chain proteins.
amino acids 232-263

FIGURE 33

GAGACACGCGAGCGGGGAGACCTCCAAGGCAGCGAGGCATCGGACATGTGTCAGCACATCTGG
GGCGCACATCCGTCGAGCCCGAGGGGAGATTTGCCGGAACAATTCAAACTGCGATATTGATCT
TGGGGGTGACTGTCCCTGGCCGGCTGTCGGGTGGGAGTGCGAGTGTGCACTCGCTCGGAAGTG
TGTGCGAGTGTGTATGTGTGTGCCGTGTCGGGCTCCCCCCTTCCCCCCGTTTTCCCGTCGA
GTGATGCACTTGGAATGAGAATCAGAGG**ATG**GAAATAGTCTGGGAGGTGCTTTTTCTTCTTCA
AGCCAATTTCATCGTCTGCATATCAGCTCAACAGAATTCACCAAAAATCCATGAAGGCTGGTG
GGCATACAAGGAGGTGGTCCAGGGAAGCTTTGTTCCAGTTCCTTCTTTCTGGGGATTGGTGAA
CTCAGCTTGGAATCTTTGCTCTGTGGGGAAACGGCAGTCGCCAGTCAACATAGAGACCAGTCA
CATGATCTTCGACCCCTTTCTGACACCTCTTCGCATCAACACGGGGGGCAGGAAGGTCAGTGG
GACCATGTACAACACTGGAAGACACGTATCCCTTCGCCTGGACAAGGAGCACTTGGTCAACAT
ATCTGGAGGGCCCATGACATACAGCCACCGGCTGGAGGAGATCCGACTACACTTTGGGAGTGA
GGACAGCCAAGGGTCGGAGCACCTCCTCAATGGACAGGCCTTCTCTGGGGAGGTGCAGCTCAT
CCACTATAACCATGAGCTATATACGAATGTCACAGAAGCTGCAAAGAGTCCAAATGGATTGGT
GGTAGTTTCTATATTTATAAAGTTTCTGATTCATCAAACCCATTTCTTAATCGAATGCTCAA
CAGAGATACTATCACAAGAATAACATATAAAAATGATGCATATTTACTACAGGGGCTTAATAT
AGAGGAACTATATCCAGAGACCTCTAGTTTCATCACTTACGATGGGTCGATGACTATCCCACC
CTGCTATGAGACAGCAAGTTGGATCATAATGAACAAACCTGTCTATATAACCAGGATGCAGAT
GCATTCCTTGCGCCTGCTCAGCCAGAACCAGCCATCTCAGATCTTTCTGAGCATGAGTGACAA
CTTCAGGCCTGTCCAGCCACTCAACAACCGCTGCATCCGCACCAATATCAACTTCAGTTTACA
GGGGAAGGACTGTCCAAACAACCGAGCCCAGAAGCTTCAGTATAGAGTAAATGAATGGCTCCT
CAAG**TAG**GGAACAAAGCCAAGAAGAATCCCACCTCAGTGAAATGCTACAACTGTGAATTGACG
TAACCTAGAATGTCCCCCTTCTTGCTTCTCTCTCCTTCTTTCCCCCAAGCCTCATTCATTCTT
GGGATTGGCCCTTTCTTCATGAAAAGTGTCTGCGAAACCATGGCAGAGGAATACATCTCTCAC
ACATACTCACAAACACACACACAAGCACTTGCACATACATACAAACACATGCAAACATACCTA
CACACACACACTCTCTTACAACCTCCATCATGGGAAGTCAAGTTTCAGAAACAAAAGTCTCAT
TCATAAGAGGTCTTAGAAGAAAATAACCAGTTAACCTGATTTCAATTTTGATACCGTTTTCCT
GAACTAATAAATCTACCCAATGAGACTTTTCAGCCTTTGTACATACAAAATTCTTCCAAAAGA
GAGAGGAGAAAATACAGCTCTGATGGCATCAAACGGACTTTGCATCAAGTAATTTCAGATAGT
GTCCTAGGATCCTTTGAGGGTGCTGGTAGCAGGTGAGCAGGACAAAGTTGACCAAGGACACTT
ATTTCTAGATTATGATTCTTCTGTTTACTCAACAATTTACAAAGAAAAAAAGGACAGACATTG
AAGAGCTACACATTGTATATATATCACCACAGACTATAAGGAAATGGAATTATTTCCCTCTTT
GTCACATATCTGTAGTAGGATTTGCCAAGATCAGAAATGATCCATTTGCTGTTTCTTGTTTTC
CAAAGGTCATACATTGTGTTTGGTTATTGTTACCAGCTCAATAAATGTGTTTAACGAGTTAAT
TTCATTTTTCTGGCTTTGGTCTGTTCTCCTTCCTTACAGGCTAAGCCCTGGCTCCATGCAACT
GCATTCTTTGATTTCACTTGTTCCTTCATCTACATGTTTTGTTCATTTGCAGCCAGTTTTTAC
TGAGTTTGTGGCAATCAGGAATGCATTTGCTAAGCAAGTATGACTTTAATTCCACTCCATGGC
TCAATCATTCACATGAGGTGAGCTTCAGCCTGAGATAGCAGGCGACAGACTTCTTGCGTTTCA
AAACTGCCATGCCCCCTGTGATGCTCCCGTGAAGGAATGCACTTTGCCTTGTAAGTTCCTGG
GAAAGGGGTATGTTTTCTCTCCAGGTGCAGCCAGATCTCACAAAGTACAAAACGAATGCCTTT
CTTTTCTTGTTTATAATGGTCACTCACTGTGTTTGGTTACTGTCAAGAAATCAATAAATGTGT
TTAACAAGTTA

FIGURE 34

MEIVWEVLFLLQANFIVCISAQQNSPKIHEGWWAYKEVVQGSFVPVPSFWGLVNSAWNLCSVG
KRQSPVNIETSHMIFDPFLTPLRINTGGRKVSGTMYNTGRHVSLRLDKEHLVNISGGPMTYSH
RLEEIRLHFGSEDSQGSEHLLNGQAFSGEVQLIHYNHELYTNVTEAAKSPNGLVVVSIFIKVS
DSSNPFLNRMLNRDTITRITYKNDAYLLQGLNIEELYPETSSFITYDGSMTIPPCYETASWII
MNKPVYITRMQMHSLRLLSQNQPSQIFLSMSDNFRPVQPLNNRCIRTNINFSLQGKDCPNNRA
QKLQYRVNEWLLK

Important features:
Signal peptide:
amino acids 1-20

Eukaryotic-type carbonic anhydrases proteins.
amino acids 126-162, 220-269, 43-91

N-glycosylation sites.
amino acids 116-119, 168-171, 302-305

FIGURE 35

```
GTCGGAACCCCCTCAGGCCACCCTCGGGAGTCCTGGGGTCCAGAGGGGTGTCCCTGTACCCCTTGCAC
ACAGGACCCTCACTCTGCAGGGATAAGCCAGCTGCGCCTGCAGCCTAGGGTGCCAAGGAGGCTGCTGA
TTGTGGCCCACAGCCTCATCTGAACGCCAGGAGACCAGGATACCGAGGCACCGGATCCCCTCTCTGTG
CCCTGGGGAGCCCCAGTGCTGCCCAGTCACCCCAGGGCTGAGGTCTGCGTCCCTAGTGGTGCAAGGCC
TGGTAGGACCACGGGGCAGGGAATGTGAGCGCCATCCGAGCTCACGGTGTCCTGAGTCGCGGCTTCGT
GACTTTGGCAGGGGCCTCCGGACCAGTGACCCCAGTCAAACCCAGAGGGTCTTGGGCGGCAGCGACGA
AGGAGGTATTCAGGCTCCAGGCCAGGTGGGGCCGGACGCCCCCAGCCATCCACC**ATG**GTGGTGGCACA
CCCCACCGCCACTGCCACCACCACGCCCACTGCCACTGTCACGGCCACCGTTGTGATGACCACGGCCA
CCATGGACCTGCGGGACTGGCTGTTCCTCTGCTACGGGCTCATCGCCTTCCTGACGGAGGTCATCGAC
AGCACCACCTGCCCCTCGGTGTGCCGCTGCGACAACGGCTTCATCTACTGCAACGACCGGGGACTCAC
ATCCATCCCCGCAGATATCCCTGATGACGCCACCACCCTCTACCTGCAGAACAACCAGATCAACAACG
CCGGCATCCCCCAGGACCTCAAGACCAAGGTCAACGTGCAGGTCATCTACCTATACGAGAATGACCTG
GATGAGTTCCCCATCAACCTGCCCCGCTCCCTCCGGGAGCTGCACCTGCAGGACAACAATGTGCGCAC
CATTGCCAGGGACTCGCTGGCCCGCATCCCGCTGCTGGAGAAGCTGCACCTGGATGACAACTCCGTGT
CCACCGTCAGCATTGAGGAGGACGCCTTCGCCGACAGCAAACAGCTCAAGCTGCTCTTCCTGAGCCGG
AACCACCTGAGCAGCATCCCCTCGGGGCTGCCGCACACGCTGGAGGAGCTGCGGCTGGATGACAACCG
CATCTCCACCATCCCGCTGCATGCCTTCAAGGGCCTCAACAGCCTGCGGCGCCTGGTGCTGGACGGTA
ACCTGCTGGCCAACCAGCGCATCGCCGACGACACCTTCAGCCGCCTACAGAACCTCACAGAGCTCTCG
CTGGTGCGCAATTCGCTGGCCGCGCCACCCCTCAACCTGCCCAGCGCCCACCTGCAGAAGCTCTACCT
GCAGGACAATGCCATCAGCCACATCCCCTACAACACGCTGGCCAAGATGCGTGAGCTGGAGCGGCTGG
ACCTGTCCAACAACAACCTGACCACGCTGCCCCGCGGCCTGTTCGACGACCTGGGGAACCTGGCCCAG
CTGCTGCTCAGGACAACCCTTGGTTTTGTGGCTGCAACCTCATGTGGCTGCGGGACTGGGTGAAGGC
ACGGGCGGCCGTGGTCAACGTGCGGGGCCTCATGTGCCAGGGCCCTGAGAAGGTCCGGGGCATGGCCA
TCAAGGACATTACCAGCGAGATGGACGAGTGTTTTGAGACGGGGCCGCAGGGCGGCGTGGCCAATGCG
GCTGCCAAGACCACGGCCAGCAACCACGCCTCTGCCACCACGCCCCAGGGTTCCCTGTTTACCCTCAA
GGCCAAAAGGCCAGGGCTGCGCCTCCCCGACTCCAACATTGACTACCCCATGGCCACGGGTGATGGCG
CCAAGACCCTGGCCATCCACGTGAAGGCCCTGACGGCAGACTCCATCCGCATCACGTGGAAGGCCACG
CTCCCCGCCTCCTCTTTCCGGCTCAGTTGGCTGCGCCTGGGCCACAGCCCAGCCGTGGGCTCCATCAC
GGAGACCTTGGTGCAGGGGGACAAGACAGAGTACCTGCTGACAGCCCTGGAGCCCAAGTCCACCTACA
TCATCTGCATGGTCACCATGGAGACCAGCAATGCCTATGTAGCTGATGAGACACCCGTGTGTGCCAAG
GCAGAGACAGCCGACAGCTATGGCCCTACCACCACACTCAACCAGGAGCAGAACGCTGGCCCCATGGC
GAGCCTGCCCCTGGCGGGCATCATCGGCGGGGCAGTGGCTCTGGTCTTCCTCTTCCTGGTCCTGGGGG
CCATCTGCTGGTACGTGCACCAGGCTGGCGAGCTGCTGACCCGGGAGAGGGCCTACAACCGGGGCAGC
AGGAAAAAGGATGACTATATGGAGTCAGGGACCAAGAAGGATAACTCCATCCTGGAAATCCGCGGCCC
TGGGCTGCAGATGCTGCCCATCAACCCGTACCGCGCCAAAGAGGAGTACGTGGTCCACACTATCTTCC
CCTCCAACGGCAGCAGCCTCTGCAAGGCCACACACACCATTGGCTACGGCACCACGCGGGGCTACCGG
GACGGCGGCATCCCCGACATAGACTACTCCTACACA**TGA**TGCCCGCCCACCCGGGCTGCCCCGCCTCA
GCCCCAGCTGCCCTGGCGTGGCCATGTGGCTTTGCCCAGCCTGCTGCAATCCAAGAGAGCAAGGAAGA
GAAATTCCATGGGTGACTTTCCTCCGCAGAAAGCAAAGTTTGGGGAGGGCTGACGATTTTGTAGAACA
CAACAGTGACAATTTTTTTTAAAAGAATAGAAGGCAGGAGGGGGAATTCGACATTGTTGAAGACATAA
TTTATACCAAGTTATGCCAGTTGGGGAGGGAAGGACTAAAAATAATATTGCAGGCAGGGCTGGGTTGG
GTTTTTTTTTTTTCCCCCCTGAACTGGAAGGATACTACCTGTACAACATCTGTGGACACCTCATGCTCT
GTTCAAGGCCATCACAAAGGAACCGCCAGGGAGAAGCAGCCGGCTCTCAAAGCTCCCACGCAGCTCTC
CCGCCACTGGCCACTCGCTGGCGACCCGATGGAAGGTTTTCAGGCTCCTCACAAAGGAGAGAGGGAAG
AAAAGATCTTTTGCCCTGGAGATATGGTCCTGAAATCTCTCCCCTGGCTTATTCCATACCATTTCCCT
TGCAGATTTGCAGAACATGGCATCTTTCACTGCATTCTTTGAACAATCATGTAGTCGATTAAAAAAA
AAAACAAACTTTTTTTTCCTAGGCTGAAGCCCTCTTCAGTTCCATGCACCACGCTCCGTAGAAGCCCC
GGCGGAAGCCGTAGCTTTCCCTGCCACCTGGAGGTGCATCTGTCTGCCTGTCTATCCCTGTCGCGGTG
TCTCTAAGTACAGATGGGTAGATAGAGCCACATGCACGGTCCTTACCGTTCTTCTTGGGTCAGTTCTT
ACCATTTCCTGAACAATAGAATTGTGAAAGTGTTAAAAA
```

FIGURE 36

MVVAHPTATATTTPTATVTATVVMTTATMDLRDWLFLCYGLIAFLTEVIDSTTCPSVCRCDNG

FIYCNDRGLTSIPADIPDDATTLYLQNNQINNAGIPQDLKTKVNVQVIYLYENDLDEFPINLP

RSLRELHLQDNNVRTIARDSLARIPLLEKLHLDDNSVSTVSIEEDAFADSKQLKLLFLSRNHL

SSIPSGLPHTLEELRLDDNRISTIPLHAFKGLNSLRRLVLDGNLLANQRIADDTFSRLQNLTE

LSLVRNSLAAPPLNLPSAHLQKLYLQDNAISHIPYNTLAKMRELERLDLSNNNLTTLPRGLFD

DLGNLAQLLLRNNPWFCGCNLMWLRDWVKARAAVVNVRGLMCQGPEKVRGMAIKDITSEMDEC

FETGPQGGVANAAAKTTASNHASATTPQGSLFTLKAKRPGLRLPDSNIDYPMATGDGAKTLAI

HVKALTADSIRITWKATLPASSFRLSWLRLGHSPAVGSITETLVQGDKTEYLLTALEPKSTYI

ICMVTMETSNAYVADETPVCAKAETADSYGPTTTLNQEQNAGPMASLPLAGIIGGAVALVFLF

LVLGAICWYVHQAGELLTRERAYNRGSRKKDDYMESGTKKDNSILEIRGPGLQMLPINPYRAK

EEYVVHTIFPSNGSSLCKATHTIGYGTTRGYRDGGIPDIDYSYT

Important features of the protein:

Transmembrane domain:

amino acids 552-573

N-glycosylation sites.

amino acids 249-252, 305-308, 642-645

Leucine zipper pattern.

amino acids 182-203, 299-320

Phospholipase A2 aspartic acid active site.

amino acids 57-67

FIGURE 37

```
GGTGACTGAAGCGAGCCTGGCCTCTTGCATCCTCCGCCTGTGTACCTCCCTCCCCTTTTTTTCCGCCT
TCTGCCAGCAGAAGCAGCAGCCGCAGCACCTGAGCCGCTACTGCCGCTCACTCAGGACAACGCTATGG
CTGAGCCTGGGCACAGCCACCATCTCTCCGCCAGAGTCAGGAGAAGAACTGAGAGGCGCATACCCCGG
CTGTGGCGGCTGCTGCTCTGGGCTGGGACCGCCTTCCAGGTGACCCAGGGAACGGGACCGGAGCTTCA
TGCCTGCAAAGAGTCTGAGTACCACTATGAGTACACGGCGTGTGACAGCACGGGTTCCAGGTGGAGGG
TCGCCGTGCCGCATACCCCGGGCCTGTGCACCAGCCTGTCTGACCCCGTCAAGGGCACCGAGTGCTCC
TTCTCCTGCAACGCCGGGGAGTTTCTGGATATGAAGGACCAGTCATGTAAGCCATGCGCTGAGGGCCG
CTACTCCCTCGGCACAGGCATTCGGTTTGATGAGTGGGATGAGCTGCCCCATGGCTTTGCCAGCCTCT
CAGCCAACATGGAGCTGGATGACAGTGCTGCTGAGTCCACCGGGAACTGTACTTCGTCCAAGTGGGTT
CCCCGGGGCGACTACATCGCCTCCAACACGGACGAATGCACAGCCACACTGATGTACGCCGTCAACCT
GAAGCAATCTGGCACCGTTAACTTCGAATACTACTATCCAGACTCCAGCATCATCTTTGAGTTTTTCG
TTCAGAATGACCAGTGCCAGCCCAATGCAGATGACTCCAGGTGGATGAAGACCACAGAGAAAGGATGG
GAATTCCACAGTGTGGAGCTAAATCGAGGCAATAATGTCCTCTATTGGAGAACCACAGCCTTCTCAGT
ATGGACCAAAGTACCCAAGCCTGTGCTGGTGAGAAACATTGCCATAACAGGGGTGGCCTACACTTCAG
AATGCTTCCCCTGCAAACCTGGCACGTATGCAGACAAGCAGGGCTCCTCTTTCTGCAAACTTTGCCCA
GCCAACTCTTATTCAAATAAAGGAGAAACTTCTTGCCACCAGTGTGACCCTGACAAATACTCAGAGAA
AGGATCTTCTTCCTGTAACGTGCGCCCAGCTTGCACAGACAAAGATTATTTCTACACACACGGCCT
GCGATGCCAACGGAGAGACACAACTCATGTACAAATGGGCCAAGCCGAAAATCTGTAGCGAGGACCTT
GAGGGGGCAGTGAAGCTGCCTGCCTCTGGTGTGAAGACCCACTGCCCACCCTGCAACCCAGGCTTCTT
CAAAACCAACAACAGCACCTGCCAGCCCTGCCCATATGGTTCCTACTCCAATGGCTCAGACTGTACCC
GCTGCCCTGCAGGGACTGAACCTGCTGTGGGATTTGAATACAAATGGTGGAACACGCTGCCCACAAAC
ATGGAAACGACCGTTCTCAGTGGGATCAACTTCGAGTACAAGGGCATGACAGGCTGGGAGGTGGCTGG
TGATCACATTTACACAGCTGCTGGAGCCTCAGACAATGACTTCATGATTCTCACTCTGGTTGTGCCAG
GATTTAGACCTCCGCAGTCGGTGATGGCAGACACAGAGAATAAAGAGGTGGCCAGAATCACATTTGTC
TTTGAGACCCTCTGTTCTGTGAACTGTGAGCTCTACTTCATGGTGGGTGTGAATTCTAGGACCAACAC
TCCTGTGGAGACGTGGAAAGGTTCCAAAGGCAAACAGTCCTATACCTACATCATTGAGGAGAACACTA
CCACGAGCTTCACCTGGGCCTTCCAGAGGACCACTTTTCATGAGGCAAGCAGGAAGTACACCAATGAC
GTTGCCAAGATCTACTCCATCAATGTCACCAATGTTATGAATGGCGTGGCCTCCTACTGCCGTCCCTG
TGCCCTAGAAGCCTCTGATGTGGGCTCCTCCTGCACCTCTTGTCCTGCTGGTTACTATATTGACCGAG
ATTCAGGAACCTGCCACTCCTGCCCCCCTAACACAATTCTGAAAGCCCACCAGCCTTATGGTGTCCAG
GCCTGTGTGCCCTGTGGTCCAGGGACCAAGAACAACAAGATCCACTCTCTGTGCTACAATGATTGCAC
CTTCTCACGCAACACTCCAACCAGGACTTTCAACTACAACTTCTCCGCTTTGGCAAACACCGTCACTC
TTGCTGGAGGGCCAAGCTTCACTTCCAAAGGGTTGAAATACTTCCATCACTTTACCCTCAGTCTCTGT
GGAAACCAGGGTAGGAAAATGTCTGTGTGCACCGACAATGTCACTGACCTCCGGATTCCTGAGGGTGA
GTCAGGGTTCTCCAAATCTATCACAGCCTACGTCTGCCAGGCAGTCATCATCCCCCCAGAGGTGACAG
GCTACAAGGCCGGGGTTTCCTCACAGCCTGTCAGCCTTGCTGATCGACTTATTGGGGTGACAACAGAT
ATGACTCTGGATGGAATCACCTCCCCAGCTGAACTTTTCCACCTGGAGTCCTTGGGAATACCGGACGT
GATCTTCTTTTATAGGTCCAATGATGTGACCCAGTCCTGCAGTTCTGGGAGATCAACCACCATCCGCG
TCAGGTGCAGTCCACAGAAAACTGTCCCTGGAAGTTTGCTGCTGCCAGGAACGTGCTCAGATGGGACC
TGTGATGGCTGCAACTTCCACTTCCTGTGGGAGAGCGCGGCTGCTTGCCCGCTCTGCTCAGTGGCTGA
CTACCATGCTATCGTCAGCAGCTGTGTGGCTGGGATCCAGANGACTACTTACGTGTGNCGAGAACCCA
AGCTATGCTCTGGTGGCATTTCTCTGCCTGAGCAGAGAGTCACCATCTGCAAAACCATAGATTTCTGG
CTGAAAGTGGGCATCTCTGCAGGCACCTGTACTGCCATCCTGCTCACCGTCTTGACCTGCTACTTTTG
GAAAAAGAATCAAAAACTAGAGTACAAGTACTCCAAGCTGGTGATGAATGCTACTCTCAAGGACTGTG
ACCTGCCAGCAGCTGACAGCTGCGCCATCATGGAAGGCGAGGATGTAGAGGACGACCTCATCTTTACC
AGCAAGAAGTCACTTTTTGGGAAGATCAAATCATTTACCTCCAAGAGGACTCCTGATGGATTTGACTC
AGTGCCGCTGAAGACATCCTCAGGAGGCCCAGACATGGACCTGTGAGAGGCACTGCCTGCCTCACCTG
CCTCCTCACCTTGCATAGCACCTTTGCAAGCCTGCGGCGATTTGGGTGCCAGCATCCTGCAACACCCA
CTGCTGGAAATCTCTTCATTGTGGCCTTATCAGATGTTTGAATTTCAGATCTTTTTTTATAGAGTACC
CAAACCCTCCTTTCTGCTTGCCTCAAACCTGCCAAATATACCCACATTTTTTTTTAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 38

MAEPGHSHHLSARVRRRTERRIPRLWRLLLWAGTAFQVTQGTGPELHACKESEYHYEYTACDS
TGSRWRVAVPHTPGLCTSLSDPVKGTECSFSCNAGEFLDMKDQSCKPCAEGRYSLGTGIRFDE
WDELPHGFASLSANMELDDSAAESTGNCTSSKWVPRGDYIASNTDECTATLMYAVNLKQSGTV
NFEYYYPDSSIIFEFFVQNDQCQPNADDSRWMKTTEKGWEFHSVELNRGNNVLYWRTTAFSVW
TKVPKPVLVRNIAITGVAYTSECFPCKPGTYADKQGSSFCKLCPANSYSNKGETSCHQCDPDK
YSEKGSSSCNVRPACTDKDYFYTHTACDANGETQLMYKWAKPKICSEDLEGAVKLPASGVKTH
CPPCNPGFFKTNNSTCQPCPYGSYSNGSDCTRCPAGTEPAVGFEYKWWNTLPTNMETTVLSGI
NFEYKGMTGWEVAGDHIYTAAGASDNDFMILTLVVPGFRPPQSVMADTENKEVARITFVFETL
CSVNCELYFMVGVNSRTNTPVETWKGSKGKQSYTYIIEENTTTSFTWAFQRTTFHEASRKYTN
DVAKIYSINVTNVMNGVASYCRPCALEASDVGSSCTSCPAGYYIDRDSGTCHSCPPNTILKAH
QPYGVQACVPCGPGTKNNKIHSLCYNDCTFSRNTPTRTFNYNFSALANTVTLAGGPSFTSKGL
KYFHHFTLSLCGNQGRKMSVCTDNVTDLRIPEGESGFSKSITAYVCQAVIIPPEVTGYKAGVS
SQPVSLADRLIGVTTDMTLDGITSPAELFHLESLGIPDVIFFYRSNDVTQSCSSGRSTTIRVR
CSPQKTVPGSLLLPGTCSDGTCDGCNFHFLWESAAACPLCSVADYHAIVSSCVAGIQXTTYVX
REPKLCSGGISLPEQRVTICKTIDFWLKVGISAGTCTAILLTVLTCYFWKKNQKLEYKYSKLV
MNATLKDCDLPAADSCAIMEGEDVEDDLIFTSKKSLFGKIKSFTSKRTPDGFDSVPLKTSSGG
PDMDL

Important features of the protein:

N-glycosylation sites:

amino acids 153-156, 390-393, 391-394, 404-407, 544-547, 576-579, 672-675, 717-720, 947-950

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

amino acids 15-18, 563-566, 709-712

Casein kinase II phosphorylation sites:

amino acids 42-45, 59-62, 81-84, 146-149, 168-171, 282-285, 331-334, 340-343, 431-434, 449-452, 465-468, 523-526, 557-560, 761-764, 780-783, 835-838, 860-863, 893-896, 949-952

Tyrosine kinase phosphorylation sites:

amino acids 50-56, 109-116

N-myristoylation sites:

amino acids 77-82, 88-93, 152-157, 268-273, 288-293, 320-325, 400-405, 405-410, 414-419, 463-468, 599-604, 616-621, 634-639, 644-649, 839-844, 874-879, 912-917, 916-921

Amidation site:

amino acids 707-710

Cell attachment sequence:

amino acids 162-164

FIGURE 39

```
GGGAAGGGGTTCTGGGCTGCCGCAGGCACACAGGCCAGAGCTTCGTGGATACCTGCAGGGCCC
AAAGGTCCCTCCCTGTTTTGAAGAGTGAGTGATGGCTATGAGGTAGCGGCCAGGCTGATCACC
CCTGCGTTGGCTGGAGGCAGAATTCTGTAAATCCTCGCCAAGTCTTTCTCCAGGCCACTGGTT
AGCTCATCTCAGCCTCCTCTGGGAGCATCAACACCAACATGGCACAGGGGACTGCAGTGGTGT
GCTTTGGACCTGTGTACCCACCCAAGGCTAAAGGCAGAGCCAGGTGACTTTGCGGGGGTCTCT
TCTCTAGGATTATCTGTACTTCCCCTCTGTCCTCTTTTACTACGGGAGATCGAGCTAGCTATA
ACCCACCTTCTTTCATGAGAACCACACTAAATTGCAAAAATTATCCCAGTGCTGGAGGAGGGC
AGCAGGTTGAGATTATGTTGGCAGGAAGAATGTTGGCATTGATTGGCACGCAGGGGACGAGAG
CTGCTTTGTGCTTTAAAGGAGCCAAGTTACACCCTGTTTAACCCTGCCTTCAAAGGGACGACT
CTGTAAGATTCTCTGCTACTTATTCAAGTTGACACG**ATG**CCCTTCACACTCCACCTGAGGTCC
CGCCTTCCCTCTGCCATAAGGAGTTTGATTCTACAAAAGAAACCAAACATCAGAAATACATCC
AGCATGGCTGGAGAGCTCCGACCAGCCAGCCTGGTGGTCCTGCCCAGGTCCCTTGCTCCAGCT
TTTGAAAGATTCTGCCAGGTCAACACTGGTCCTCTACCCCTGCTGGGCCAGAGTGAGCCAGAA
AAGTGGATGCTGCCCCCTCAAGGTGCTATCTCAGAGACCAGGATGGGCCATCCCCAGTTCTGG
AAATACGAGTTCGGTGCCTGCACCGGTAGCCTGGCTTCGCTGGAGCAGTACTCGGAGCAGCTG
AAGGACATGGTGGCCTTCTTCCTGGGCTGCAGCTTCTCCCTGGAGGAGGCCTTGGAGAAAGCG
GGGCTCCCCAGAAGAGACCCAGCAGGTCACAGCCAGGCGGGTGCATACAAGACAACAGTGCCT
TGTGTTACCCATGCTGGCTTCTGCTGCCCTCTGGTGGTCACGATGAGGCCCATTCCCAAGGAC
AAGCTGGAAGGGCTGGTGCGGGCCTGCTGCTCCCTCGGAGGTGAGCAGGGGCAACCTGTTCAC
ATGGGCGACCCAGAACTGTTGGGAATCAAAGAGCTTTCCAAACCTGCCTACGGGGATGCCATG
GTGTGTCCCCCAGGGGAGGTTCCAGTGTTCTGGCCTTCTCCGCTGACCAGTCTCGGAGCTGTC
AGCAGCTGTGAGACCCCACTGGCTTTTGCCAGCATCCCAGGCTGCACAGTTATGACTGACCTG
AAGGATGCAAAGGCTCCACCTGGTTGTCTCACCCCAGAGAGAATTCCAGAGGTCCATCACATT
TCCCAAGATCCTCTGCACTACAGCATCGCGTCAGTCTCTGCTTCTCAGAAGATCAGAGAACTA
GAGTCTATGATCGGCATAGACCCAGGGAACCGGGGGATTGGGCACCTGCTCTGTAAAGATGAG
CTGCTGAAGGCCTCTCTCTCGCTGTCCCATGCCCGCTCAGTGCTCATCACCACTGGGTTCCCC
ACACATTTCAATCATGAGCCTCCAGAAGAGACAGATGGCCCACCAGGAGCTGTTGCTCTGGTT
GCCTTCCTGCAGGCCTTGGAGAAGGAGGTCGCCATAATCGTTGACCAGAGAGCCTGGAACTTG
CACCAGAAGATTGTTGAAGATGCTGTTGAGCAAGGTGTTCTGAAGACGCAGATCCCGATATTA
ACTTACCAAGGTGGATCAGTGGAAGCTGCTCAGGCATTCCTGTGCAAAAATGGGGACCCGCAG
ACACCTAGATTTGACCACCTGGTGGCCATAGAGCGTGCCGGAAGAGCTGCTGATGGCAATTAC
TACAATGCAAGGAAGATGAACATCAAGCACTTGGTTGACCCCATTGACGATCTTTTTCTTGCT
GCGAAGAAGATTCCTGGAATCTCATCAACTGGAGTCGGTGATGGAGGCAACGAGCTTGGGATG
GGTAAAGTCAAGGAGGCTGTGAGGAGGCACATACGGCACGGGGATGTCATCGCCTGCGACGTG
GAGGCTGACTTTGCCGTCATTGCTGGTGTTTCTAACTGGGGAGGCTATGCCCTGGCCTGCGCA
CTCTACATCCTGTACTCATGTGCTGTCCACAGTCAGTACCTGAGGAAAGCAGTCGGACCCTCC
AGGGCACCTGGAGATCAGGCCTGGACTCAGGCCCTCCCGTCGGTCATTAAGGAAGAAAAAATG
CTGGGCATCTTGGTGCAGCACAAAGTCCGGAGTGGCGTCTCGGGCATCGTGGGCATGGAGGTG
GATGGGCTGCCCTTCCACAACACCCACGCCGAGATGATCCAGAAGCTGGTGGACGTCACCACG
GCACAGGTG**TAA**CCGTCCATGTTCCGTGTGAGCAGAGTCCCTACCAACGGGCAGGTCTGCATC
CGGGGAGAATGCAGCTGCTTCTGGCGACAATCCTGCTAGTAAACACTGGTCTTCGGTGAGCAA
CGAACACTCGCCTGGCCTGGGAAACTGCATGCCCACTTTCTGGGAGGGGTTAGTGCAGGTGCC
GTGGACAAAGGACAACATTTCTCTGGGGCTTTTTAACTTTTATTCCTAAGACTCTAAAGGCGT
TGATTTCAACCCTCCTTCACTCTGGCTTCTTCAGGCAACCCACGTGGTCTCCTATGAGAATCT
TCTCGACAGTTACTTATGGGGACACTTGTGAACAATTAACTGCCAGGGCAGAGCATGAGAACA
AACATTCCCAGGCCATGTAGGATAGGATACTCCAGACTCCAGTCATCCTCCCCATCCATGGT
TTCTGTTACTCATGGTTTCAGTTACTCATAGCCAACTGCAGACCGAAAATACTAAATGAAAAA
TTTCAGAAATAAACAACTCTTAAGTTTTAAAAAAAAA
```

FIGURE 40

MPFTLHLRSRLPSAIRSLILQKKPNIRNTSSMAGELRPASLVVLPRSLAPAFERFCQVNTGPL
PLLGQSEPEKWMLPPQGAISETRMGHPQFWKYEFGACTGSLASLEQYSEQLKDMVAFFLGCSF
SLEEALEKAGLPRRDPAGHSQAGAYKTTVPCVTHAGFCCPLVVTMRPIPKDKLEGLVRACCSL
GGEQGQPVHMGDPELLGIKELSKPAYGDAMVCPPGEVPVFWPSPLTSLGAVSSCETPLAFASI
PGCTVMTDLKDAKAPPGCLTPERIPEVHHISQDPLHYSIASVSASQKIRELESMIGIDPGNRG
IGHLLCKDELLKASLSLSHARSVLITTGFPTHFNHEPPEETDGPPGAVALVAFLQALEKEVAI
IVDQRAWNLHQKIVEDAVEQGVLKTQIPILTYQGGSVEAAQAFLCKNGDPQTPRFDHLVAIER
AGRAADGNYYNARKMNIKHLVDPIDDLFLAAKKIPGISSTGVGDGGNELGMGKVKEAVRRHIR
HGDVIACDVEADFAVIAGVSNWGGYALACALYILYSCAVHSQYLRKAVGPSRAPGDQAWTQAL
PSVIKEEKMLGILVQHKVRSGVSGIVGMEVDGLPFHNTHAEMIQKLVDVTTAQV

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 358-378, 517-539

N-glycosylation site.

amino acids 28-32

Tyrosine kinase phosphorylation site.

amino acids 444-452

N-myristoylation site.

amino acids 98-104, 102-108, 123-129, 149-155, 181-187, 190-196,
238-244, 308-314, 399-405, 413-419, 448-454, 477-483, 482-488,
487-493

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 233-244, 531-542

FIGURE 41

CTTTCCTGTTTTATCCGCAGCCCTTTTCTTCTTTGAGTTAGTAAAGATTTATTCTGTAACCTG

ACACTCATCTGGCCCTTTGCAGTTTGCCAGCCATATTCCCATGTGATTTCCCACTGGATCCAG

GCCCCCATCCGGCTGGCAGGAGGGGGCTCTGACGTACAGGTTGGAAATCAGAAGTCTGTGAGA

GCGCGGGAGTGCATGGCAGCTCTGGGTCCCAGACCTGGCCCGACCCCTCTGCTTCACCTCCAG

CTCTGCTGCTCCTCTACTCTTGGGTCGAGATCCCTTTGGAGCCACAGCGAGGAACCCTGTGGT

CCTCAGGCAGGTGTACCTTGAGTCAGCCAGGAGCCCTCTTTTCCTGTGTCAAAGCCTGCCCTC

GGGCTCTGCTCACCTCTGGTGACCCTCCAAGATGCCCCTGCCCTCAGTTTCCCCTCATGATCT

GGCCTCTGCCCCCTTCTCTAGCCACAGCCTCTAGTACACTTTAGCAATACCACCAGACTAGTT

AGAGTTCCCCACTCACCAAGCAAGAC**ATG**CAGTTTCATGCCTCTGTGCCTTCGCTCATGCTGT

TTCTTCCGACTGGAATGCCTTCCCCTGCTCCTCCTGCCTTGTCTGCCTGGCAAGTTCATCTCT

CACGATCCCCTCAAAGGCCCCCTCCTCCAGGAAGGCAACCCCTGTGCCCCTCCCCTCCAGGCT

ACCTCTGCACTTTGTCAATGCTTCTCTTGTGGCACTTATCACACTGTATTTTACTTGTTTACA

TGTTTGTCTCCCCTTCTAGACTG**TGA**ATCCTTAAGGGCATGGACTGTATCTTATGCATCTCTG

TATTTCTGCGCCTAGCACGGTGCCTAGCACACAGTAGGCGCTCAATAAATGTTGAATGAATGA

ATGATTT

FIGURE 42

MQFHASVPSLMLFLPTGMPSPAPPALSAWQVHLSRSPQRPPPPGRQPLCPSPPGYLCTLSMLL
LWHLSHCILLVYMFVSPSRL

Important features of the protein:

Signal peptide:

amino acids 1-22

Microbodies C-terminal targeting signal.

amino acids 81-83

FIGURE 43

GTTTCCAACAAGGATGATATGAAGACTTCCCTGAAGAAAGTTGTGAAGGGACCTCCTACGAG**A**

TGATGATGCAGTGTGTGTCCCGCATGTTGGCCCACCCCCTGCATGTCATCTCAATGCGCTGCA

TGGTCCAGTTTGTGGGACGGGAGGCCAAGTACAGTGGTGTGCTGAGCTCCATTGGGAAGATTT

TCAAAGAGGAAGGGCTGCTGGGATTCTTCGTTGGATTAATCCCTCACCTCCTGGGCGATGTGG

TTTTCTTGTGGGGCTGTAACCTGCTGGCCCACTTCATCAATGCCTACCTGGTGGATGACAGCT

TCAGCCAGGCCCTGGCCATCCGGAGCTATACCAAGTTCGTGATGGGGATTGCAGTGAGCATGC

TGACCTACCCCTTCCTGCTAGTTGGCGACCTCATGGCTGTGAACAACTGCGGGCTGCAAGCTG

GGCTCCCCCTTACTCCCCAGTGTTCAAATCCTGGATTCACTGCTGGAAGTACCTGAGTGTGC

AGGGCCAGCTCTTCCGAGGCTCCAGCCTGCTTTTCCGCCGGGTGTCATCAGGATCATGCTTTG

CCCTGGAG**TAA**CCTGAATCATCTAAAAAACACGGTCTCAACCTGGCCACTGTGGGTGAGGCCT

GACCACCTTGGGACACCTGCAAGACGACTCCAACCCAACAACAACCAGATGTGCTCCAGCCCA

GCCGGGCTTCAGTTCCATATTTGCCATGTGTCTGTCCAGATGTGGGGTTGAGCGGGGGTGGGG

CTGCACCCAGTGGATTGGGTCACCCGGCAGACCTAGGGAAGGTGAGGCGAGGTGGGGAGTTGG

CAGAATCCCCATACCTCGCAGATTTGCTGAGTCTGTCTTGTGCAGAGGGCCAGAGAATGGCTT

ATGGGGGCCCAGGTTGGATGGGGAAAGGCTAATGGGGTCAGACCCCACCCCGTCTACCCCTCC

AGTCAGCCCAGCGCCCATCCTGCAGCTCAGCTGGGAGCATCATTCTCCTGCTTTGTACATAGG

GTGTGGTCCCCTGGCACGTGGCCACCATCATGTCTAGGCCTATGCTAGGAGGCAAATGGCCAG

GCTCTGCCTGTGTTTTTCTCAACACTACTTTTCTGATATGAGGGCAGCACCTGCCTCTGAATG

GGAAATCATGCAACTACTCAGAATGTGTCCTCCTCATCTAATGCTCATCTGTTTAATGGTGAT

GCCTCGCGTACAGGATCTGGTTACCTGTGCAGTTGTGAATACCCAGAGGTTGGGCAGATCAGT

GTCTCTAGTCCTACCCAGTTTTAAAGTTCATGGTAAGATTTGACCTCATCTCCCGCAAATAAA

TGTATTGGTGATTTGGAAAAAAAAAAAAAAAAA

FIGURE 44

MMMQCVSRMLAHPLHVISMRCMVQFVGREAKYSGVLSSIGKIFKEEGLLGFFVGLIPHLLGDV
VFLWGCNLLAHFINAYLVDDSFSQALAIRSYTKFVMGIAVSMLTYPFLLVGDLMAVNNCGLQA
GLPPYSPVFKSWIHCWKYLSVQGQLFRGSSLLFRRVSSGSCFALE

Important features of the protein:

Signal peptide:

amino acids 1-18

Transmembrane domains:

amino acids 51-72, 97-114

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 160-163

N-myristoylation sites.

amino acids 34-39, 100-105, 123-128, 165-170

FIGURE 45

```
GCTCACTCTTTGGGTCCACACTGCCTTTATGAGCTGTAACACTCACTGGGAATGTCTGCAGCT
TCACTCCTGAAGCCAGCGAGACCACGAACCCACCAGGAGGAACAAACAACTCCAGACGCGCAG
CCTTAAGAGCTGTAACACTCACCGCGAAGGTCTGCAGCTTCACTCCTGAGCCAGCCAGACCAC
GAACCCACCAGAAGGAAGAAACTCCAAACACATCCGAACATCAGAAGGAGCAAACTCGTGACA
CGCCACCTTTAAGAACCGTGACACTCAACGCTAGGGTCCGCGGCTTCATTCTTGAAGTCAGTG
AGACCAAGAACCCACCAATTCCGGACACGGCAAAGTAACATCCTAGAC**ATG**GCTTTAGAGATC
CACATGTCAGACCCCATGTGCCTCATCGAGAACTTTAATGAGCAGCTGAAGGTTAATCAGGAA
GCTTTGGAGATCCTGTCTGCCATTACGCAACCTGTAGTTGTGGTAGCGATTGTGGGCCTCTAT
CGCACTGGCAAATCCTACCTGATGAACAAGCTGGCTGGGAAGAACAAGGGCTTCTCTGTTGCA
TCTACGGTGCAGTCTCACACCAAGGGAATTTGGATATGGTGTGTGCCTCATCCCAACTGGCCA
AATCACACATTAGTTCTGCTTGACACCGAGGGCCTGGGAGATGTAGAGAAGGCTGACAACAAG
AATGATATCCAGATCTTTGCACTGGCACTCTTACTGAGCAGCACCTTTGTGTACAATACTGTG
AACAAAATTGATCAGGGTGCTATCGACCTACTGCACAATGTGACAGAACTGACAGATCTGCTC
AAGGCAAGAAACTCACCTGACCTTGACAGGGTTGAAGATCCTGCTGACTCTGCGAGCTTCTTC
CCAGACTTAGTGTGGACTCTGAGAGATTTCTGCTTAGGCCTGGAAATAGATGGGCAACTTGTC
ACACCAGATGAATACCTGGAGAATTCCCTAAGGCCAAAGCAAGGTAGTGATCAAAGAGTTCAA
AATTTCAATTTGCCCCGTCTGTGTATACAGAAGTTCTTTCCAAAAAAGAAATGCTTTATCTTT
GACTTACCTGCTCACCAAAAAAAGCTTGCCCAACTTGAAACACTGCCTGATGATGAGCTAGAG
CCTGAATTTGTGCAACAAGTGACAGAATTCTGTTCCTACATCTTTAGCCATTCTATGACCAAG
ACTCTTCCAGGTGGCATCATGGTCAATGGATCTCGTCTAAAGAACCTGGTGCTGACCTATGTC
AATGCCATCAGCAGTGGGGATCTGCCTTGCATAGAGAATGCAGTCCTGGCCTTGGCTCAGAGA
GAGAACTCAGCTGCAGTGCAAAAGGCCATTGCCCACTATGACCAGCAAATGGGCCAGAAAGTG
CAGCTGCCCATGGAAACCCTCCAGGAGCTGCTGGACCTGCACAGGACCAGTGAGAGGGAGGCC
ATTGAAGTCTTCATGAAAAACTCTTTCAAGGATGTAGACCAAAGTTTCCAGAAGAATTGGAG
ACTCTACTAGATGCAAAACAGAATGACATTTGTAAACGGAACCTGGAAGCATCCTCGGATTAT
TGCTCGGCTTTACTTAAGGATATTTTTGGTCCTCTAGAAGAAGCAGTGAAGCAGGGAATTTAT
TCTAAGCCAGGAGGCCATAATCTCTTCATTCAGAAAACAGAAGAACTGAAGGCAAAGTACTAT
CGGGAGCCTCGGAAAGGAATACAGGCTGAAGAAGTTCTGCAGAAATATTTAAAGTCCAAGGAG
TCTGTGAGTCATGCAATATTACAGACTGACCAGGCTCTCACAGAGACGGAAAAAAAGAAGAAA
GAGGCACAAGTGAAAGCAGAAGCTGAAAAGGCTGAAGCGCAAAGGTTGGCGGCGATTCAAAGG
CAGAACGAGCAAATGATGCAGGAGAGGGAGAGACTCCATCAGGAACAAGTGAGACAAATGGAG
ATAGCCAAACAAAATTGGCTGGCAGAGCAACAGAAATGCAGGAACAACAGATGCAGGAACAG
GCTGCACAGCTCAGCACAACATTCCAAGCTCAAAATAGAAGCCTTCTCAGTGAGCTCCAGCAC
GCCCAGAGGGCTGTTAATAACGATGATCCATGTGTTTTACTC**TAA**AGTGCTAAATATGGGAGT
TTCCTTTTTTTACTCTTTGTCACTGATGACACAACAGAAAGAAACTGTAGACCTTGGGACAA
TCAACATTTAAATAAACTTTATAATTATTAAA
```

FIGURE 46

MALEIHMSDPMCLIENFNEQLKVNQEALEILSAITQPVVVVAIVGLYRTGKSYLMNKLAGKNK
GFSVASTVQSHTKGIWIWCVPHPNWPNHTLVLLDTEGLGDVEKADNKNDIQIFALALLLSSTF
VYNTVNKIDQGAIDLLHNVTELTDLLKARNSPDLDRVEDPADSASFFPDLVWTLRDFCLGLEI
DGQLVTPDEYLENSLRPKQGSDQRVQNFNLPRLCIQKFFPKKKCFIFDLPAHQKKLAQLETLP
DDELEPEFVQQVTEFCSYIFSHSMTKTLPGGIMVNGSRLKNLVLTYVNAISSGDLPCIENAVL
ALAQRENSAAVQKAIAHYDQQMGQKVQLPMETLQELLDLHRTSEREAIEVFMKNSFKDVDQSF
QKELETLLDAKQNDICKRNLEASSDYCSALLKDIFGPLEEAVKQGIYSKPGGHNLFIQKTEEL
KAKYYREPRKGIQAEEVLQKYLKSKESVSHAILQTDQALTETEKKKKEAQVKAEAEKAEAQRL
AAIQRQNEQMMQERERLHQEQVRQMEIAKQNWLAEQQKMQEQQMQEQAAQLSTTFQAQNRSLL
SELQHAQRAVNNDDPCVLL

Important features of the protein:

Transmembrane domains:

amino acids 31-49, 114-131

N-glycosylation sites.

amino acids 90-94, 144-148, 287-291, 563-567

N-myristoylation sites.

amino acids 45-51, 283-289

Prenyl group binding site.

amino acids 583-588

ATP/GTP-binding site motif A (P-loop).

amino acids 45-53

FIGURE 47

CACTCATTCATTCCAAAGGGTCTCTCAAGGCAATGGTAATGTGCAAGGAGGTGATACCTAAAT
GAATGACCAAAAGAACATGCTTCTGCTTTTGTGTGTCTCCTACATTTTAGACATTTGTTTGTT
TCTCTTGGTAGCCTTTAAATTCCTTGAAGCCCAGGACCATGTCTCACTTACCTTTGTGTTTCC
ACTAACTAGTCTACCTCCTGGAATTGGCAGATACTCAGTGAAAGCCTGTGAAATAAGTGATGT
CTATTTCTAGCATATTATTCTGAGATTTAATGATAGATTTAGTGATTGAATGAGATTTCCATT
TTCAAATACAGCAAAAGCATAACTATTTTCATTCATTCATATTCATTCAACTTCATTCTCAAA
ATTAGGTCCTGAGTTAACTAATAATTACCTTTGAAATGTGTGGGTTATTTGAGGCAATCAGGT
GGTGACATTGAGCTCTCAGCCAGAGTTTGTTTCTGGAATTGATTCAGTTCCATTGCATTGATT
TTTGTTCTCAGAAGCCAAGGTTTCCCATGAAAAATCATTCCCACTTGAATTGGGCTGTGATTC
TTGCTGCGTTTAAGTAAAGGAAGCCTCTTGGTTCTAGTTCTGCAAACTTACACACTGAACTGG
GACAAGTTTTTGTTTAGAGTAATGGCTGGGAAAAGAGGAACCTTTCATTTTATTCAGAAGTCA
AAAACAAAGGCCTCCCAGCCACCTGGAGATGTTTTGTTGCAGACACCAGCCTGGCTCTGTCTT
TATGCCTAACAATTGAGCATCCAGTCTTCTTTGTGCTGGGACCATTGCTCAGCTCTGCAAGGG
GAAAAGAGGGAGAAAGCCAGAGCTGCCAGGCTTCTTGCACTGGGGCCGGGGGAGGGTTCCTGG
GAAGCAGGTGCTCTCTGGCTTCTTGGTACGTGAGGCTCTCGGAGCTGCCTCTCCTCTGACCCT
CAGGTCCTCACCGAGTTTGCTCCAGGAGTATATTGAAAACATACCCAGTGCTCTCTCAAGCAC
CCACTGCTTAGAGGGCCCAGATTTCTTTTCCTTCTTTCCCTTGCAGAGCTGGAGACTGCATCG
GGCATCTGGTGTTTAAACTAAACAGGAAACTGACTAAAGGTCCACAGTGCTCATTGTGTAGA
CTAGCTGCCCTCCG**ATG**GGTGCTCTGATTATCAGTGGTTCCAGTGCAGGGCCTGTCACTAAAC
AGGCCTCACTTCCTCCTTGGGGGCTTTCCCATGGGAGGTGTGGCTTTTTACTCTACATGGAAA
TGACTCTCTGCAGCCACAGAACACAGTCATTTTCTGAATTATCCCAGTCTCTCATGCGCCCTG
GATTCCTCCAGATGCCTTATATCTCTTGTGCAAAGTTGTCTAAAATTTGGTTCCCAGCTTCCA
AGCCTTGCCTTTTGGCCTTCCTGGAAGTATTTTTGTTGATGAGTCGTCTGTCATTATTCTCTA
AAATGATTTGCTTTTTGTTTCTTTCATTCCTATTTCCACCCCACATATACACACATGCTTCT**T**
AACTTAGGGGATTACATGCCAATAAATCTATTGTTGAAAATGCACTAATACTATCGCAAAGAC
GAAAATTCACAGGCTGAACCGTTGTAAGTCCATATGCTCCTCAACTTACATGTGTGATGGAGT
TATGCCCAAATAAGTCCATCGTCAAGTTGAAAAATCAAAATCAAGCCATCTTAGGTTGAGGAC
CATTTGTTTGTACCTCCAAAGATGTCATATCTTTAAACATACTCCCTAGCTTTTCTTTTTACT
TTTTATTTTGAAGTAATTATAGAATCACAGAAAGTTGCAAAAAA

FIGURE 48

MGALIISGSSAGPVTKQASLPPWGLSHGRCGFLLYMEMTLCSHRTQSFSELSQSLMRPGFLQM
PYISCAKLSKIWFPASKPCLLAFLEVFLLMSRLSLFSKMICFLFLSFLFPPHIYTHAS

Important features of the protein:

Signal peptide:

amino acids 1-41

Transmembrane domain:

amino acids 88-107

Casein kinase II phosphorylation site.

amino acids 47-50

N-myristoylation site.

amino acids 24-29

FIGURE 49

GGCTTCTACAGTCCACAACACCCACCAGCCCCAGGCCCAGCAGAATGAGCCCAGTGAGTGCCGGGGCTCCCAGTT
TGGCTGTTGCTATGACAACGTGGCCACTGCAGCCGGTCCTCTTGGGGAAGGCTGTGTGGGCCAGCCCAGCCATGC
CTACCCCGTGCGGTGCCTGCTGCCCAGTGCCCATGGCTCTTGTGCAGACTGGGCTGCCCGCTGGTACTTCGTTGC
CTCTGTGGGCCAATGTAACCGCTTCTGGTATGGCGGCTGCCATGGCAATGCCAATAACTTTGCCTCGGAGCAAGA
GTGCATGAGCAGCTGCCAGGGATCTCTCC**ATG**GGCCCCGTCGTCCCCAGCCTGGGGCTTCTGGAAGGAGCACCCA
CACGGATGGTGGCGGCAGCAGTCCTGCAGGCGAGCAGGAACCCAGCCAGCACAGGACAGGGGCCGCGGTGCAGAG
AAAGCCCTGGCCTTCTGGTGGTCTCTGGCGGCAAGACCAACAGCCTGGGCCAGGGGAGGCCCCCCACACCCAGGC
CTTTGGAGAATGGCCATGGGGGCAGGAGCTTGGGTCCAGGGCCCCTGGACTGGGTGGAGATGCCGGATCACCAGC
GCCACCCTTCCACAGCTCCTCCTACAGATCTCACTTCCCACCTCTCCAGGATTAGCTTGGCAGGTGTGGAGCCCT
CGTTGGTGCAGGCAGCCCTGGGGCAGTTGGTGCGGCTCTCCTGCTCAGACGACACTGCCCCGGAATCCCAGGCTG
CCTGGCAGAAAGATGGCCAGCCCATCTCCTCTGACAGGCACAGGCTGCAGTTCGACGGATCCCTGATCATCCACC
CCCTGCAGGCAGAGGACGCGGGCACCTACAGCTGTGGCAGCACCCGGCCAGGCCGCGACTCCCAGAAGATCCAAC
TCCGCATTATAGGGGGTGACATGGCCGTGCTGTCTGAGGCTGAGCTGAGCCGCTTCCCTCAGCCCAGGGACCCAG
CTCAGGACTTTGGCCAAGCGGGGGCTGCTGGGCCCCTGGGGGCCATCCCCTCTTCACACCCACAGCCTGCAAACA
GGCTGCGTTTGGACCAGAACCAGCCCCGGGTGGTGGATGCCAGTCCAGGCCAGCGGATCCGGATGACCTGCCGTG
CCGAAGGCTTCCCGCCCCCAGCCATCGAGTGGCAGAGAGATGGGCAGCCTGTCTCTTCTCCCAGACACCAGCTGC
AGCCTGATGGCTCCCTGGTCATTAGCCGAGTGGCTGTGAAGATGGCGGCTTCTACACCTGTGTCGCTTTCAATG
GGCAGGACCGAGACCAGCGATGGGTCCAGCTCAGAGTTCTGGGGGAGCTGACAATCTCAGGACTGCCCCCTACTG
TGACAGTGCCAGAGGGTGATACGGCCAGGCTATTGTGTGTGGTAGCAGGAGAAAGTGTGAACATCAGGTGGTCCA
GGAACGGGCTACCTGTGCAGGCTGATGGCCACCGTGTCCACCAGTCCCCAGATGGCACGCTGCTCATTTACAACT
TGCGGGCCAGGGATGAGGGCTCCTACATGTGCAGTGCCTACCAGGGGAGCCAGGCAGTCAGCCGCAGCACCGAGG
TGAAGGTGGTCTCACCAGCACCCACCGCCCAGCCCAGGGACCCTGGCAGGGACTGCGTCGACCAGCCAGAGCTGG
CCAACTGTGATTTGATCCTGCAGGCCCAGCTTTGTGGCAATGAGTATTACTCCAGCTTCTGCTGTGCCAGCTGTT
CACGTTTCCAGCCTCACGCTCAGCCCATCTGGCAG**TAG**GGATGAAGGCTAGTTCCAGCCCCAGTCCAAAATAGTT
CATAGGGCTAGGGAGAAAGGAAGATGGACTCTTGGCTTCCTCTCTCTGGCTGGCAAAGGGAGTTATCTTCTGGAA
TACATTAGCTCTTTCAAAAACCCACCCAGTGTTTAGCCTCAACGGCAGCCAGTTACCAGCTTCTCTCTGTAGCCT
TCAGCAGTGTTTGCATCTCTGACATAACCACAGGCTGCTGTTTTCAAGAAGAGCAATCTGTTTGGATAAGAAAAA
CCTTTACTTTACAGCTTCCCTTTATAATTTGTTACACAGGAATAGTTAAATGCATTTGTTTGTTTGTTTTTTGAG
ACGGAGTTTCACTCTTGTTGCCCAGGCTGGAGGGCAATGGCGCGATCTCAGCTCACTGCAACCTCCGTCTCCTGG
GTTCTTGATTCTCCTGTGTCAGCCTTCTGAGTAGCTGGGATTACAGATGCCTATCACCATGCCTGGGTAATTTTT
GTATTTTTAGTTGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAACTTCTGACCTCAGATGATCTGCCC
GCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCACGCCCAGCCATCAATGCATTTTTTTTATTTT
TTTTTTGAGACAGAGTTTCGCACTTCTTGCCCAGGCTGGAGTACAATGGTGCGATCTTGGCTCACTGCAACCTCC
ACCTCCTGGGTTCAAGCGCTTCTCCAGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGTATGTGCCACCATGCCT
GGCTAATTTTGTATTTTTGGTGGAGACGGGGTTTCTCCATGTTGGTCAGACTGGTCTTGAACTCCCGACCTCAGG
TAATCCGCCCGCCTCCGCCTCCCAAAATGCTGGGATTAGAGGTGTGAGCCACTGTGCCCAGCCCATCAATGTGTT
TTAAAGCTAGCTGTCAGGGTTCCACTTAATTTAAAGCTGGGCAGGGAGATGTGTAATGATTTCAAAGTTAACACC
TGTTTGTTTTCTAAAGGGCATGCCAAGTCCTGCTGTATCAGGGAAGTATTCTGTGCTAAAATCAGCGATGGTTCA
TTGCTCTAGTCTCTCTCACCCTTCTAGGCAGTGCATCAGTCAGCTCTAAATCTGGTGCAGAGGGTTAACAGCATA
ACCCTTGTTGGCAAAATGGAATAGATGTTAAGACCTCAAATAGGGATTTGGGATGAAACAGCTGCAGTTAGCACT
GTTATCTGAGCATGAAAGAACTGGAAACGCTCCTTACGTCGAGATGTTGGACCTTGAAGCCCTCCTGAGGCCAAC
ATGCAAATCTGGCTGTGACGGTTCATCTGACACCTGTGTAAAGCTGACCAGCCTGCTCTGTACAGTGACAATGAG
GAGCCCCTCTCTTCCTTAAGTAGGAATCTGTGAAGCAAAATGTTTGCTGCCAAAGACAAATCAGACTGTCAGTCA
TTAAAAACAGCATTAGCAGGATGAGGATAGCAATGGGGAAGGGTTGTGGGCAATGCAGTAACAGGGAAATGGCTT
CAGAAATGGTTTGAGTTGGAAGACAACATTCTTCATCTCTCAGGACTTCTAATTCCTTGATGCTAAAAGAAGAGG
CATGGATTCTATGAGCTTCCAAGTCCCTTTCCACTTTAACCTTCTACAAATCTTTCAGAGGACTGCCTAGTAGCA
AAGGTTATTCCTGGACACAGGAAAGACGGGCATTACAGGGACCAAAGCTCTGAAAGGTGACTTTTATTACCAACA
CACTGGCTGGAAAAGGGACAAACCACATCACGGGTGAGTGATACTTCTCAGTCTTCTCTACTCATTCAACAAAGG
AAATGTGGGCTGGGGCAGAGGTCTTTTTTCATTTAATACTGGAAAAATATTGAAGAGCATCCATGTTCACTTATG
GCTGGTTTTGCTATAGAAATTGGAAAATAAAGGCCACTTTTTTG

FIGURE 50

MGPVVPSLGLLEGAPTRMVAAAVLQASRNPASTGQGPRCRESPGLLVVSGGKTNSLGQGRPPT
PRPLENGHGGRSLGPGPLDWVEMPDHQRHPSTAPPTDLTSHLSRISLAGVEPSLVQAALGQLV
RLSCSDDTAPESQAAWQKDGQPISSDRHRLQFDGSLIIHPLQAEDAGTYSCGSTRPGRDSQKI
QLRIIGGDMAVLSEAELSRFPQPRDPAQDFGQAGAAGPLGAIPSSHPQPANRLRLDQNQPRVV
DASPGQRIRMTCRAEGFPPPAIEWQRDGQPVSSPRHQLQPDGSLVISRVAVEDGGFYTCVAFN
GQDRDQRWVQLRVLGELTISGLPPTVTVPEGDTARLLCVVAGESVNIRWSRNGLPVQADGHRV
HQSPDGTLLIYNLRARDEGSYMCSAYQGSQAVSRSTEVKVVSPAPTAQPRDPGRDCVDQPELA
NCDLILQAQLCGNEYYSSFCCASCSRFQPHAQPIWQ

Important features of the protein:

Signal peptide:

amino acids 1-16

Tyrosine kinase phosphorylation site.

amino acids 392-400

N-myristoylation sites.

amino acids 9-15, 50-56, 112-118, 146-152, 173-179, 195-201, 220-226, 229-235, 280-286, 306-312, 336-342, 397-403

Myelin P0 protein.

amino acids 153-182

FIGURE 51

```
CAGGCAGAAGCGAACAAAGACCCAGCAAGAGAAGGCAGAGGCTAAGACCCATCCCGTATCTGC
TCTCCTGAAATAATTCTGGAGTCATGCCTGAAATGCCAGAGGACATGGAGCAGGAGGAAGTTA
ACATCCCTAATAGGAGGGTTCTGGTTACTGGTGCCACTGGGCTTCTTGGCAGAGCTGTACACA
AAGAATTTCAGCAGAATAATTGGCATGCAGTTGGCTGTGGTTTCAGAAGAGCAAGACCAAAAT
TTGAACAGGTTAATCTGTTGGATTCTAATGCAGTTCATCACATCATTCATGATTTTCAGCCCC
ATGTTATAGTACATTGTGCAGCAGAGAGAAGACCAGATGTTGTAGAAAATCAGCCAGATGCTG
CCTCTCAACTTAATGTGGATGCTTCTGGGAATTTAGCAAAGGAAGCAGCTGCTGTTGGAGCAT
TTCTCATCTACATTAGCTCAGATTATGTATTTGATGGAACAAATCCACCTTACAGAGAGGAAG
ACATACCAGCTCCCCTAAATTTGTATGGCAAACAAAATTAGATGGAGAAAGGCTGTCCTGG
AGAACAATCTAGGAGCTGCTGTTTTGAGGATTCCTATTCTGTATGGGGAAGTTGAAAAGCTCG
AAGAAAGTGCTGTGACTGTTATGTTTGATAAAGTGCAGTTCAGCAACAAGTCAGCAAACATGG
ATCACTGGCAGCAGAGGTTCCCCACACATGTCAAAGATGTGGCCACTGTGTGCCGGCAGCTAG
CAGAGAAGAGAATGCTGGATCCATCAATTAAGGGAACCTTTCACTGGTCTGGCAATGAACAGA
TGACTAAGTATGAAATGGCATGTGCAATTGCAGATGCCTTCAACCTCCCCAGCAGTCACTTAA
GACCTATTACTGACAGCCCTGTCCTAGGAGCACAACGTCCGAGAAATGCTCAGCTTGACTGCT
CCAAATTGGAGACCTTGGGCATTGGCCAACGAACACCATTTCGAATTGGAATCAAAGAATCAC
TTTGGCCTTTCCTCATTGACAAGAGATGGAGACAAACGGTCTTTCATTAGTTTATTTGTGTTG
GGTTCTTTTTTTTTTAAATGAAAGTATAGTATGTGGCACTTTTTAAAGAACAAAGGAAATA
GTTTTGTATGAGTACTTTAATTGTGACTCTTAGGATCTTTCAGGTAAATGATGCTCTTGCACT
AGTGAAATTGTCTAAAGAAACTAAAGGGCAGTCATGCCCTGTTTGCAGTAATTTTTCTTTTTA
TCATTTTGTTTGTCCTGGCTAAACTTGGAGTTTGAGTATAGTAAATTATGATCCTTAAATATT
TGAGAGTCAGGATGAAGCAGATCTGCTGTAGACTTTTCAGATGAAATTGTTCATTCTCGTAAC
CTCCATATTTTCAGGATTTTTGAAGCTGTTGACCTTTTCATGTTGATTATTTTAAATTGTGTG
AAATAGTATAAAAATCATTGGTGTTCATTATTTGCTTTGCCTGAGCTCAGATCAAAATGTTTG
AAGAAAGGAACTTTATTTTTGCAAGTTACGTACAGTTTTTATGCTTGAGATATTTCAACATGT
TATGTATATTGGAACTTCTACAGCTTGATGCCTCCTGCTTTTATAGCAGTTTATGGGGAGCAC
TTGAAAGAGCGTGTGTACATGTATTTTTTTTCTAGGCAAACATTGAATGCAAACGTGTATTTT
TTTAATATAAATATATAACTGTCCTTTTCATCCCATGTTGCCGCTAAGTGATATTTCATATGT
GTGGTTATACTCATAATAATGGGCCTTGTAAGTCTTTTCACCATTCATGAATAATAATAAATA
TGTACTGCTGGCATGTAATGCTTAGTTTTCTTGTATTTACTTCTTTTTTTAAATGTAAGGACC
AAACTTCTAAACTAATTGTTCTTTTGTTGCTTTAATTTTTAAAAATTACATTCTTCTGATGTA
ACATGTGATACATACAAAGAATATAGTTTAATATGTATTGAAATAAAACACAATAAAATT
```

FIGURE 52

MPEMPEDMEQEEVNIPNRRVLVTGATGLLGRAVHKEFQQNNWHAVGCGFRRARPKFEQVNLLD
SNAVHHIIHDFQPHVIVHCAAERRPDVVENQPDAASQLNVDASGNLAKEAAAVGAFLIYISSD
YVFDGTNPPYREEDIPAPLNLYGKTKLDGEKAVLENNLGAAVLRIPILYGEVEKLEESAVTVM
FDKVQFSNKSANMDHWQQRFPTHVKDVATVCRQLAEKRMLDPSIKGTFHWSGNEQMTKYEMAC
AIADAFNLPSSHLRPITDSPVLGAQRPRNAQLDCSKLETLGIGQRTPFRIGIKESLWPFLIDK
RWRQTVFH

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 105-127

N-glycosylation site.

amino acids 197-201

N-myristoylation site.

amino acids 303-309

Short-chain dehydrogenases/reductases family proteins.

amino acids 18-30

FIGURE 53

```
TGGGCTCCCTCCAGCACTGCTGTTGCCTGCTGCCTAAGATGGGTGACACTTGGGCCCAGCTTCCCTGGCCTGGGC
CACCCCACCCAGCAATGCTGCTGATCTCCCTCCTCTTGGCAGCCGGGTTGATGCACTCGGATGCCGGCACCAGCT
GCCCCGTCCTTTGCACATGCCGTAACCAGGTGGTGGATTGTAGCAGCCAGCGGCTATTCTCCGTGCCCCCAGACC
TGCCAATGGACACCCGAAACCTCAGCCTGGCCCACAACCGCATCACAGCAGTGCCGCCTGGCTACCTCACATGCT
ACATGGAGCTCCAGGTGCTGGATTTGCACAACAACTCCTTAATGGAGCTGCCCCGGGGCCTCTTCCTCCATGCCA
AGCGCTTGGCACACTTGGACCTGAGCTACAACAATTTCAGCCATGTGCCAGCCGACATGTTCCAGGAGGCCCATG
GGCTAGTCCACATCGACCTGAGCCACAACCCCTGGCTGCGGAGGGTGCATCCCCAGGCCTTTCAGGGCCTCATGC
AGCTCCGAGACCTGGACCTCAGTTATGGGGGCCTGGCCTTCCTCAGCCTGGAGGCTCTTGAGGGCCTACCGGGGC
TGGTGACCCTGCAGATCGGTGGCAATCCCTGGGTGTGTGGCTGCACCATGGAACCCCTGCTGAAGTGGCTGCGAA
ACCGGATCCAGCGCTGTACAGCAGATTCTCAGCTGGCTGAGTGCCGGGGCCCTCCTGAAGTCGAGGGCGCCCCGC
TCTTCTCACTCACTGAGGAGAGCTTCAAGGCCTGCCACCTGACCCTGACCCTGGATGATTACCTATTCATTGCGT
TCGTGGGCTTCGTGGTCTCCATTGCTTCTGTGGCCACCAACTTCCTCCTGGGCATCACTGCCAACTGCTGCCACC
GCTGGAGCAAGGCCAGTGAAGAGGAAGAGATCTGACATGCCTGCCTCTCATCCCTCCATGCTGCTGACCGCCACA
GCTGCTGGCCACCAGACGCCCTCCCTGATTGCTCACTCTGGTTCCATGGTGACCTGGCTGCCTCAGTCATGGTTC
AAGCAAGGTGGGGACACTCATTTTGTATGAGCATCTGCTTTGGGCCAGGCGGCACGCTAGGAATTGGGAACATCA
GATGAACTGACTCAGTCCCTGCCCTCAAGGCACTTCCCTCTGGTCAAGGAGAGAGATCCAAAAACTATTCCCTTT
AAGACTATATGTCAGGACTCTGAGCACGTCATTATGGAGGCCCAGAGGAGGAGCCATCATCTGTATCTAGCAATG
TCCATGAGAATTATAAGATTAGAGTGATTTGTGAACTGGGTCATCAGGAAATATCTACTTTGTCAGGTAGGCAAA
GAAGGGTGTCTGCACATGGCAGAGGCCAGAATATGCATAGTGTGCTGTGTTGAGAAGAGTGAACAGTTCCTGGTC
ACTTACTTGTATAGAGGGGGTGTGGCACAGAACTCAAACCTACCCCTCACCTCCTGACACCAAAACTGTCAGCTC
TCAGCAATGCCAGCCACTGCCTACAGGGAGTAAGAACACCTCTATGACAGCCCCTGGCCTCCTTCCACCAGCAGC
TACCAGGTGAGACCACCTCCCAGTGACTGCCCCCATATGACCAAATGTCACCAGTTGGTGAGGTCCCAGGCAGCA
GGCTGAGGATGGACACTTTCAATGCCCTTGCTCCTGCCTCTCACTCAAGTTTTGCTTCAGAAGAGAGAGGCAGGA
GGCCCAGCAACTGGGGCAGCAAGAGTCCTGGCACCTTGGGATCCTAATCATGTGACTGTTCTTGCCACAGTGCTC
ATGCCACAGGGTCTCACCAGGAAAGTGCACTGTGGGCCACAGACCCACAGCCTGGCAGCACCCAGAGCTAAAAGG
GGACAAAGGCAGCACAGTTATGACCATATGAGGCTTTGCATTTTCTTCTAAGCAACTTACCCACGTTAAGCATGA
GGGTGAGAGAGCTATTAAATACTAAGCCCTTGCCAGTGTCAGGTACTTTGAAAAGCTCTCTGCACAAACCATTCC
CTTTGACACACACACACACAAATCTTTTGAGGTGAACGCTGTTGTTCCCATTTTACGGATGAGGCAACTAAGGCT
CAGAGAGGTTAAAGTCACATGCCACTATGAGCAAGATAAAGTCTGTGCTCTTTCTACTGCCCCATCCAAGTTGGG
GAACATCACCATTCCCTCTAGAGTTATATAAATTCAAATTCAACTAGAGCTGACAAAGTTCCTCATAAGGTCCAG
GCACTCCTCTGGGCACTTTTATATCTATTGACTCACTTCTTTCAATTCTCACAGCAACACTGCCTGGTGGTTTTT
ATTATCCCCATTTGACAGATGAATTAATCGTAGAGAGTTGAGTGACTTACCCAAGGTTGTCTGGATAAGCCCTAG
AAGGAAGGCGGTAGGCAGCTCCATTCAGGGAAACTGCATCTAATCAGTCAGTCAAAAATCAAGTAACTTTACGAG
CAAAGCACAATTATCATCATCGTGGTCTTCTTCATCAGTTTCGTCAGCAGCATCATTATCTTCCCTCTATTTGTT
CAGCACCGGATAGTTCATGAGTATTTTTGCATCATTCTCCTTGACTTTTCACATCCCTGTGCAGGAGGTAAATCA
AACATCAGTAATCCTGTTTTACAGATGGGGAAAAAAGTCTCAAGGTTGGATATGACTTGCTATGTGGCAAGGTTG
GGGCTCAACCCTAACACAGTTCTCTTTCCAGTGCTTTCTCAAGTGCTTGGGGAAGAGAATGCCTCAGAAGGCTGG
GTAGTGGGGCCCTGGAATTCAGCATCCATGAATGTGCTAGTGGATAAGCTAAATAGAAGGCAGCCAAACCCATCT
GCTGTACAGATTGAACTATGCTCACGGTAGGGCAAATTGCAGGCTCTGAAACAGAGACTACACAGGTAACACCTG
AATAGGAGACTCCTGCTTTACAATGTGTAGATAAAACATCAGCAATGGTGGCCATGGTGGCAGTCATGTGAAAAG
TAAGATCTTTGGGAATCAAGAAAGGAAGCTGTGTTAACCACTCCTGCTCAAGCCCTGCTGCGTGTGTTGCAAGAG
ATACTAAGAGAGCAAGAAAGCTATAGGTGAGAACCTCTGCAGTTTAGGAGAAGAACATCAAGGCACAGTCCAACA
TGCTGATAAGTCTGGCCAGGAGGAGAATTAAAACAGGGCTTTCCACACCTCCCTTGCCCCAAGCTCCAGCGGTA
TTCTATCAGCCCATCCTCCTGGAAAGCCTGAAAGGAATGAAGGAGGCTAATAAGTCATCTTCCAGGAAGGCATCC
CTCACTCGTGCTTCCCTGAGCTAGTCAACCAAAGAGTCTTCAGAAACTTTGCTAGACCTGAAGTACTTGAACCT
GTGTCCCCTGAATCTTTCTTACAACATCTGGGACAAATCCCTGGTCCTGTGACATCCGAAGCAGAACTGTGCCCT
GCTCTCTCCTTCTGTGATGACCAAGGATGGTGAACTCAAGTTGTTCTCTACAAGCCAGGCCAGCAACCTAAATAC
TTGGAGAGGAACTTTTAGAAACTATAATCCTGACAAAATAGAAAAGTTTCCCATAGGGGCATACCATAATACTAT
AATAACCTCCCAGGAACTATTGTTTGCCAAAATGTAGTTAATATATTTTAAGATATATGCTTTTTTGCATAGGAC
TAGAACCAGAAAAGACACCAAATGCCCCCTTGACATCAATGTCCTTTCTAGTGGGACAATTTGGTCTCCATTAAT
GCCAAACCTTTCTGAACAGGATACATGGCTTTTAAAGGACAGATGTTTCTCCTGCTGCTAGAAGTTCCTCAGTTT
ACTAGAGCACAATGAGGAAAGTATTCAACCTCCCTACTGCCAAGGAATTCCCTGCTTCTCCCCCACCGCCATCAT
CTTGTCCAAGCTATCAGAAGCAACCTTCTAGAGATAATCTAACAATCCTGATTAGAATTGCTCCCATATCCCTGG
TGACCACAGGCTTCATTCAAATTGTCCAAACTGGTTAACATGTATGTGATGGGGTATCTCTGCATCTGTATGTCT
GTCTGCGAGGTTCCTTGTATATTGGCTGTCCGCTGACTTGGGACAGATCTCTCTAGAACTTGGGTTCAGTTCTCT
GACATAGTCCACTCAGCCATAGGCTGAGTGGCTAAATATGCATAAATAAGCATGCCTAAATAGGCATATATAGGT
TGGTGCAAAAGTAATTGCGGTTTTTTGCCATTAAAAATGATGGCAAAAATCCCAATTACTTTTGCGTCAATCTAAT
ATTACATTGCTTGATAGATTAAGATGGAATCCCACCAGGTTTAGGGTAGGACTGGATGCTCAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA
```

FIGURE 54

MLLISLLLAAGLMHSDAGTSCPVLCTCRNQVVDCSSQRLFSVPPDLPMDTRNLSLAHNRITAV

PPGYLTCYMELQVLDLHNNSLMELPRGLFLHAKRLAHLDLSYNNFSHVPADMFQEAHGLVHID

LSHNPWLRRVHPQAFQGLMQLRDLDLSYGGLAFLSLEALEGLPGLVTLQIGGNPWVCGCTMEP

LLKWLRNRIQRCTADSQLAECRGPPEVEGAPLFSLTEESFKACHLTLTLDDYLFIAFVGFVVS

IASVATNFLLGITANCCHRWSKASEEEI

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 241-260

N-glycosylation sites.

amino acids 52-55, 81-84, 107-110

Tyrosine kinase phosphorylation site.

amino acids 148-154

N-myristoylation sites.

amino acids 11-15, 263-268

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 175-185

Leucine zipper pattern.

amino acids 77-98

FIGURE 55

GGCTGCGCCCAGGCCGGCGGGCCCAGCAGCTGCGAACCGCCGGCGCACCACCTGTTTCCGCGC
CCGGGGACTTCCCCGGCGGGGCTCAGAAGTGTGGGGTCGGTCGCTTGGCTTCCCCTGGCGTCA
GCGACCCAGGGTAACCTCCTCCACTGCTGCGTGCCGTGCAGGCCTGCCTGTGTGAGAGCCACG
TGTGCCGCGCTCTGGGCACAGCCTTGGAAAGTCAGGACCGCGACGGCAGCAGAGCAGAAACCT
TACAGAAACATGAAGCCCTCAACCATCTGCTACTCAGTTATTCGGGGCTGACGGCGGCTTCTA
GAACATCCAGGTGTTCTGCAGATGCGAGAACTCATCCTGTAGTCACCAGATGGAGTCCCAAAC
AGCCAAGCAGATGTAAGGCCTGTGCTGTGGCTCTGAGGCCCTGAATACAGAAGGGTCACTTTC
TTAGTGGCCAAAGAGCAGTTGTTGACATTGATGTCTAATTATTGAACACGACCAGTCATTTTA
CTGAGCTGCAGTGAGGAAACACTGACCATAGAAGATCAAGCCAAATGAGGGATTGCAAATTTC
CTGATTCTTTTGAATTAGGATTCCAGATGGGGGCCTCATTTCTACAGCCCCCAACATTCCTAT
AGCCGTTATCACTGCCATCACCACTGCCACCAGCATCTTCTTGCAGATTCCACCCCTGCTCCC
CAGAGACTTCCTGCTTTGAAAGTGAGCAGAAAGGAAGCTCTCAGAAAAATCTCTAGTGGTGGC
TGCCGTCGCTCCAGACAATCGGAATCCTGCCTTCACCACC**ATG**GGCTGGCTTTTTCTAAAGGT
TTTGTTGGCGGGAGTGAGTTTCTCAGGATTTCTTTATCCTCTTGTGGATTTTTGCATCAGTGG
GAAAACAAGAGGACAGAAGCCAAACTTTGTGATTATTTTGGCCGATGACATGGGGTGGGGTGA
CCTGGGAGCAAACTGGGCAGAAACAAAGGACACTGCCAACCTTGATAAGATGGCTTCGGAGGG
AATGAGGTTTGTGGATTTCCATGCAGCTGCCTCCACCTGCTCACCCTCCCGGGCTTCCTTGCT
CACCGGCCGGCTTGGCCTTCGCAATGGAGTCACACGCAACTTTGCAGTCACTTCTGTGGGAGG
CCTTCCGCTCAACGAGACCACCTTGGCAGAGGTGCTGCAGCAGGCGGGTTACGTCACTGGGAT
AATAGGCAAATGGCATCTTGGACACCACGGCTCTTATCACCCCAACTTCCGTGGTTTTGATTA
CTACTTTGGAATCCCATATAGCCATGATATGGGCTGTACTGATACTCCAGGCTACAACCACCC
TCCTTGTCCAGCGTGTCCACAGGGTGATGGACCATCAAGGAACCTTCAAAGAGACTGTTACAC
TGACGTGGCCCTCCCTCTTTATGAAAACCTCAACATTGTGGAGCAGCCGGTGAACTTGAGCAG
CCTTGCCCAGAAGTATGCTGAGAAAGCAACCCAGTTCATCCAGCGTGCAAGCACCAGCGGGAG
GCCCTTCCTGCTCTATGTGGCTCTGGCCCACATGCACGTGCCCTTACCTGTGACTCAGCTACC
AGCAGCGCCACGGGGCAGAAGCCTGTATGGTGCAGGGCTCTGGGAGATGGACAGTCTGGTGGG
CCAGATCAAGGACAAAGTTGACCACACAGTGAAGGAAAACACATTCCTCTGGTTTACAGGAGA
CAATGGCCCGTGGGCTCAGAAGTGTGAGCTAGCGGGCAGTGTGGGTCCCTTCACTGGATTTTG
GCAAACTCGTCAAGGGGGAAGTCCAGCCAAGCAGACGACCTGGGAAGGAGGGCACCGGGTCCC
AGCACTGGCTTACTGGCCTGGCAGAGTTCCAGTTAATGTCACCAGCACTGCCTTGTTAAGCGT
GCTGGACATTTTTCCAACTGTGGTAGCCCTGGCCCAGGCCAGCTTACCTCAAGGACGGCGCTT
TGATGGTGTGGACGTCTCCGAGGTGCTCTTTGGCCGGTCACAGCCTGGGCACAGGGTGCTGTT
CCACCCCAACAGCGGGGCAGCTGGAGAGTTTGGAGCCCTGCAGACTGTCCGCCTGGAGCGTTA
CAAGGCCTTCTACATTACCGGTGGAGCCAGGGCGTGTGATGGGAGCATGGTGCCTGAGCTGCA
GCATAAGTTTCCTCTGATTTTCAACCTGGAAGACGATACCGCAGAAGCTGTGCCCCTAGAAAG
AGGTGGTGCGGAGTACCAGGCTGTGCTGCCCGAGGTCAGAAAGGTTCTTGCAGACGTCCTCCA
AGACATTGCCAACGACAACATCTCCAGCGCAGATTACACTCAGGACCCTTCAGTAACTCCCTG
CTGTAATCCCTACCAAATTGCCTGCCGCTGTCAAGCCGCA**TAA**CAGACCAATTTTTATTCCAC
GAGGAGGAGTACCTGGAAATTAGGCAAGTTTGCTTCCAAATTTCATTTTTACCCTCTTTACAA
ACACACGCTTTAGTTTAGTCTTGGAGTTTAGTTTTGGAGTTAGCCTTGCATATCCCTTCTGTA
TCCTGTCCCCCTCCACGCCGACCCGAGAGCAGCTGAGCTGCGCTGGCTCTGGGCAGGGAGTG
TGCCTTAATGGGAAGCACACGGGCTTTGGAGTCAGGCACAGGTGCCAGCTCCAGCTTTTGAAC
TTGGGCAATTGTTTAACCTAACCTGCAAGTTGATTTTGAGGGTTAAATAAAGGCATACATGAA
AATGCCTGGCAACTTTAAAAAAAAAAAA

FIGURE 56

MGWLFLKVLLAGVSFSGFLYPLVDFCISGKTRGQKPNFVIILADDMGWGDLGANWAETKDTAN
LDKMASEGMRFVDFHAAASTCSPSRASLLTGRLGLRNGVTRNFAVTSVGGLPLNETTLAEVLQ
QAGYVTGIIGKWHLGHHGSYHPNFRGFDYYFGIPYSHDMGCTDTPGYNHPPCPACPQGDGPSR
NLQRDCYTDVALPLYENLNIVEQPVNLSSLAQKYAEKATQFIQRASTSGRPFLLYVALAHMHV
PLPVTQLPAAPRGRSLYGAGLWEMDSLVGQIKDKVDHTVKENTFLWFTGDNGPWAQKCELAGS
VGPFTGFWQTRQGGSPAKQTTWEGGHRVPALAYWPGRVPVNVTSTALLSVLDIFPTVVALAQA
SLPQGRRFDGVDVSEVLFGRSQPGHRVLFHPNSGAAGEFGALQTVRLERYKAFYITGGARACD
GSMVPELQHKFPLIFNLEDDTAEAVPLERGGAEYQAVLPEVRKVLADVLQDIANDNISSADYT
QDPSVTPCCNPYQIACRCQAA

Important features of the protein:
Signal peptide:
amino acids 1-16

Transmembrane domain:
amino acids 353-373

N-glycosylation sites.
amino acids 117-120, 215-218, 356-359, 397-500

N-myristoylation sites.
amino acids 12-17, 33-38, 52-57, 97-102, 101-106, 113-118, 158-163, 328-333, 388-393, 418-423, 435-440, 436-441

Amidation site.
amino acids 382-385

Sulfatases signature 2.
amino acids 129-138

FIGURE 57

TGGACAAGACACCTCCAGGAGCCCAGCTCACAGCCACCGGTACCTTCTTCCAGGACAAGCTGG
GGGCCTCCATGGGCGCCTGAGGGCCAGGCGCCAGGGCCGTGGGCACGAGT**ATG**GTGAGACACC
AGCCCCTGCAGTACTACGAGCCACAGCTGTGCCTCTCCTGCCTCACGGGCATCTACGGCTGCC
GTTGGAAGCGCTACCAGCGCTCCCATGATGATACCACACCGGGCACAGCGCCATTCCTGCATG
TGGGGGCTGTGGCAGCAGTCACCATGCTCTCCTGGATCGTGGCAGGACAGTTCGCCCGTGCAG
AGCGGACCTCCTCCCAGGTGACCATTCTCTGTACCTTCTTCACCGTGGTGTTTGCCCTCTACC
TGGCCCCTCTCACCATCTCCTCTCCCTGCATCATGGAGAAGAAAGACCTCGGCCCCAAGCCTG
CTCTCATTGGCCACCGCGGGGCCCCCATGCTGGCTCCAGAGCACACGCTCATGTCCTTCCGGA
AGGCCCTCGAGCAGAAGCTGTACGGGCTCCAGGCTGACATTACCATCAGCCTGGACGGCGTGC
CCTTCCTCATGCATGACACCACCCTGCGGCGCACCACCAACGTGGAGGAGGAGTTCCCGGAGC
TGGCCCGCAGGCCTGCCTCCATGCTTAACTGGACCACCCTGCAGAGACTCAACGCTGGCCAGT
GGTTCCTGAAGACTGACCCCTTCTGGACAGCCAGCTCCCTGTCACCCTCCGACCACAGAGAGG
CCCAGAACCAGTCCATCTGCAGCCTGGCAGAGCTCCTGGAGCTGGCCAAGGGCAATGCCACAC
TGCTGCTCAACCTGCGTGACCCGCCCCGGGAGCACCCCTACCGCAGCAGTTTTATCAACGTGA
CTCTGGAGGCCGTGCTGCACTCCGGCTTCCCCCAGCACCAGGTCATGTGGCTGCCTAGCAGGC
AGAGGCCCCTGGTGCGGAAGGTGGCTCCCGGCTTCCAACAGACATCAGGCTCCAAGGAGGCAG
TCGCCAGCCTGCGGAGAGGCCACATCCAGCGGCTGAACCTGCGCTACACTCAGGTGTCCCGCC
AGGAGCTCAGGGACTACGCGTCCTGGAACCTGAGTGTGAACCTCTACACAGTCAACGCACCGT
GGCTCTTCTCCCTGCTGTGGTGTGCGGGGGTCCCATCCGTCACCTCTGACAACTCCCACACCC
TGTCCCAGGTGCCTTCCCCCCTCTGGATCATGCCCCCGGACGAGTACTGTCTCATGTGGGTCA
CTGCCGACCTGGTCTCCTTCACCCTCATCGTGGGCATCTTCGTGCTCCAGAAGTGGCGCCTGG
GTGGCATACGGAGCTACAACCCTGAGCAGATCATGCTGAGTGCTGCGGTGCGCCGGACCAGCC
GGGACGTCAGCATCATGAAGGAGAAGCTTATTTTCTCAGAGATCAGCGATGGTGTAGAGGTCT
CCGATGTGCTCTCCGTATGTTCAGACAACAGTTATGACACATATGCCAACAGCACCGCCACCC
CTGTGGGCCCCCGAGGGGGTGGCAGCCACACCAAGACCCTCATAGAGCGGAGTGGGCGT**TAG**C
TGAAGACATGTCTGTCCCACCTGTACCTGACACAGAAGCTGGGGAGCCTAGGAGAGCTGGTGG
AAGTGTGTCTGAACTCGGAGTGCTCTGGGAGCGGGCTCCACAGCCTCCTTGTGGGCTCCAGCC
CCTTGTCAGCCGCAGCCTCTCTTGAGGGGGACTCCCTGTCTCCTGAGGCCCAGCTGGGCCAGG
ACTCCATCCTTTCAGATGCCCCTGCAGGCCTGGGGCTCCTTCTGGGAAGTATGGGGCCTAGGG
CTTGGTCCCCCTCTTCTGAGGCCCTCTCCTGTATCCCGACCTGGAAGCTTTGATGGGTCATGG
GCCATGCCATACCCCCTGTGGCAATGGAGTGTGTGGATGCTCACCTGTGCCATCTGTCCTCCT
GTCTGTGCCAGGAGGCACCTGAGTTCTCTGCTGTTATCCTGCCCCAAGGGCCTGGGCCGAGCC
TCTACCTGAAGCAACTCTGCTCTTCCTGTCAGTCTCAAAGCACAAGGAGGTTCAGCCCAGGAG
GAAGCCAGCTGCAATGTGGAGACACGTCCTCCTCCCCAACCCACCTCATGCCACCGCCAACCC
CCTGCCCCAGGAGCGGGCCTGAGCCACGTCCCCTAGGAGCAGCTGGAGATGGCCAAAAGAGTG
AGCTCAGGACTACTGGATCCCATGCCCAGGTGTCCAGCAGACCTCAAGGCAGAAGGGTCACCT
AACCCAGGAGTCCACAGACTGATGTGACCTCAGGTTCCCACATCAGTGGCCACAGGGCAGGGC
CCACCTGGTAGAAGTGTTCTGGATATGGCCAGGGTGGGTGTGTGGCTAAGTGGGCCTGAACAG
AGGGAACCTAGGGCCCTTGGCCAATGTGATTAAAGCTGCCATCTTGAAA

FIGURE 58

MVRHQPLQYYEPQLCLSCLTGIYGCRWKRYQRSHDDTTPGTAPFLHVGAVAAVTMLSWIVAGQ

FARAERTSSQVTILCTFFTVVFALYLAPLTISSPCIMEKKDLGPKPALIGHRGAPMLAPEHTL

MSFRKALEQKLYGLQADITISLDGVPFLMHDTTLRRTTNVEEEFPELARRPASMLNWTTLQRL

NAGQWFLKTDPFWTASSLSPSDHREAQNQSICSLAELLELAKGNATLLLNLRDPPREHPYRSS

FINVTLEAVLHSGFPQHQVMWLPSRQRPLVRKVAPGFQQTSGSKEAVASLRRGHIQRLNLRYT

QVSRQELRDYASWNLSVNLYTVNAPWLFSLLWCAGVPSVTSDNSHTLSQVPSPLWIMPPDEYC

LMWVTADLVSFTLIVGIFVLQKWRLGGIRSYNPEQIMLSAAVRRTSRDVSIMKEKLIFSEISD

GVEVSDVLSVCSDNSYDTYANSTATPVGPRGGGSHTKTLIERSGR

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 47-61, 77-93, 335-350, 380-399

N-glycosylation sites.

amino acids 182-186, 217-221, 233-237, 255-259, 329-333, 462-466

Tyrosine kinase phosphorylation site.

amino acids 130-139

N-myristoylation sites.

amino acids 21-27, 48-54, 294-300, 404-410, 442-448, 473-479

FIGURE 59

CCTGAGCAAACACAGCAGCCCGAGTGTTCCCAAGGCCAAA**ATG**CTGAGAACGTCCACTCCTAA
TCTGTGTGGTGGTCTGCATTGCCGGGCCCCCTGGCTCTCTTCTGGCATTCTCTGCCTCTGCCT
CATATTCTTGTTAGGCCAGGTGGGCTTGCTGCAGGGACACCCCAGTGCCTGGATTACGGGCC
CCCTTTCCAGCCCCCTCTGCACCTTGAGTTTTGCTCTGACTATGAGTCCTTCGGCTGCTGTGA
TCAGCACAAGGACCGCCGCATCGCTGCCCGGTACTGGGACATCATGGAATATTTTGATCTGAA
GAGACATGAGCTGTGTGGAGATTACATTAAAGACATCCTTTGCCAGGAGTGCTCGCCCTACGC
AGCCCACCTCTACGACGCCGAAAACACCCAGACGCCTCTCCGGAATCTCCCGGGCCTCTGCTC
TGATTACTGCTCTGCCTTCCATTCTAACTGTCACTCAGCCATTTCCCTGCTGACCAATGACCG
CGGCCTCCAGGAGTCTCATGGAAGGGACGGTACCCGCTTCTGCCACCTCCTGGACCTTCCTGA
CAAGGACTATTGCTTCCCTAATGTCCTGAGGAACGACTATCTCAACCGCCACCTGGGCATGGT
GGCCCAAGATCCTCAGGGCTGCCTGCAGCTCTGCCTGAGCGAGGTGGCCAACGGGCTGAGGAA
CCCCGTCTCCATGGTCCATGCTGGGGACGGCACCCATCGCTTCTTTGTTGCCGAGCAGGTAGG
AGTGGTGTGGGTCTACCTCCCTGATGGGAGTCGCCTGGAGCAACCCTTCCTGGACCTCAAGAA
CATCGTGTTGACCACCCCATGGATCGGGGATGAGAGAGGCTTCTTGGGGTTGGCTTTTCACCC
CAAATTCCGCCACAATCGCAAGTTCTATATTTATTATTCGTGCCTGGACAAGAAGAAGGTAGA
AAAGATCCGAATTAGTGAGATGAAGGTTTCTCGGGCTGATCCTAACAAAGCTGACCTGAAATC
AGAGAGGGTCATCTTGGAGATTGAAGAACCAGCCTCAAACCATAATGGCGGACAACTTCTTTT
TGGCCTGGATGGCTATATGTACATATTCACTGGGGACGGGGGACAGGCTGGAGATCCCTTTGG
CCTGTTTGGAAATGCTCAGAACAAAAGTTCCCTGCTGGGAAAAGTTTTAAGGATCGATGTGAA
CAGGGCAGGCTCACATGGCAAGCGGTACCGAGTCCCCTCGGACAATCCATTTGTTTCTGAGCC
AGGGGCCCACCCCGCCATCTATGCCTATGGGATCAGGAACATGTGGCGTTGTGCTGTGGACCG
AGGGGACCCCATCACGCGCCAGGGCCGAGGCCGGATATTCTGTGGGGACGTGGGCCAGAACAG
GTTTGAAGAGGTTGACCTCATTTTGAAAGGTGGAAACTATGGCTGGAGAGCAAAGGAAGGGTT
TGCATGTTATGACAAAAAACTTTGTCACAATGCCTCTTTGGATGATGTTCTGCCAATCTATGC
TTATGGCCATGCAGTGGGGAAGTCAGTCACTGGAGGTTATGTCTATCGTGGTTGTGAATCCCC
AAATCTCAATGGCCTGTATATCTTTGGAGACTTCATGAGTGGTCGACTTATGGCTTTGCAGGA
AGATAGAAAAAACAAGAAATGGAAGAAGCAGGATCTTTGCCTGGGCAGCACCACGTCCTGTGC
CTTCCCAGGGCTGATCAGCACCCATAGCAAGTTCATCATCTCCTTTGCTGAAGATGAAGCAGG
GGAGCTGTATTTCCTGGCGACCTCTTACCCAAGTGCCTATGCACCACGTGGATCTATTTACAA
GTTTGTTGACCCCTCAAGGCGAGCACCCCAGGCAAGTGCAAATACAAGCCAGTGCCCGTGAG
AACCAAGAGTAAGCGGATCCCGTTCAGACCACTCGCCAAGACAGTCTTGGACTTGCTAAAGGA
ACAATCAGAGAAAGCTGCTAGAAAATCTTCCAGTGCAACCTTAGCTTCTGGCCCAGCCCAGGG
TTTGTCTGAGAAAGGCTCCTCCAAGAAGCTGGCTTCTCCTACAAGCAGCAAGAATACATTGCG
AGGGCCTGGTACAAAGAAGAAAGCCAGAGTGGGGCCCCACGTCCGCCAGGGCAAGAGGAGGAA
GAGCCTGAAAAGCCACAGTGGCAGGATGAGGCCATCAGCAGAGCAGAAGCGAGCTGGCAGAAG
TCTCCCT**TGA**CCTATTGGTCAAGGTGGCCGACAGGGTGACGTGAGAGAGGAGAGCCACCTCAT
CAAATGAAAGTCACTGCTGAATAAAGACCTTAGAAGTCTGGGAAGCCAGGGTAGAGGTGGGGC
AGGGCGGTTTTCCTCTCCCTGGGAAATCTTGCTGTCTACTGAATAAATAAATGCACCTTCTCT
GTATGCAGTGCTTCTGTGGGAGACCATATCCCAGATTGCTGGTGCACCTGGGTTATGGTAAGC
ACTAGTCCATGAGCCTGCTTGGAATCACACTGGATGTCTCCGTTTTGTCTTGTAAATGCCTAC
AACCTGAGGTAATAAATCAACATTTGCTCA

FIGURE 60

MLRTSTPNLCGGLHCRAPWLSSGILCLCLIFLLGQVGLLQGHPQCLDYGPPFQPPLHLEFCSD
YESFGCCDQHKDRRIAARYWDIMEYFDLKRHELCGDYIKDILCQECSPYAAHLYDAENTQTPL
RNLPGLCSDYCSAFHSNCHSAISLLTNDRGLQESHGRDGTRFCHLLDLPDKDYCFPNVLRNDY
LNRHLGMVAQDPQGCLQLCLSEVANGLRNPVSMVHAGDGTHRFFVAEQVGVVWVYLPDGSRLE
QPFLDLKNIVLTTPWIGDERGFLGLAFHPKFRHNRKFYIYYSCLDKKKVEKIRISEMKVSRAD
PNKADLKSERVILEIEEPASNHNGGQLLFGLDGYMYIFTGDGGQAGDPFGLFGNAQNKSSLLG
KVLRIDVNRAGSHGKRYRVPSDNPFVSEPGAHPAIYAYGIRNMWRCAVDRGDPITRQGRGRIF
CGDVGQNRFEEVDLILKGGNYGWRAKEGFACYDKKLCHNASLDDVLPIYAYGHAVGKSVTGGY
VYRGCESPNLNGLYIFGDFMSGRLMALQEDRKNKKWKKQDLCLGSTTSCAFPGLISTHSKFII
SFAEDEAGELYFLATSYPSAYAPRGSIYKFVDPSRRAPPGKCKYKPVPVRTKSKRIPFRPLAK
TVLDLLKEQSEKAARKSSSATLASGPAQGLSEKGSSKKLASPTSSKNTLRGPGTKKKARVGPH
VRQGKRRKSLKSHSGRMRPSAEQKRAGRSLP

Important features of the protein:

Signal peptide:

amino acids 1-41

Transmembrane domain:

amino acids 17-36

N-glycosylation sites.

amino acids 372-376, 480-484

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 645-649, 699-703

Tyrosine kinase phosphorylation site.

amino acids 81-89

N-myristoylation sites.

amino acids 11-17, 37-43, 156-162, 165-171, 357-363, 365-371, 368-374, 408-414, 459-465, 548-554, 557-563

Amidation sites.

amino acids 391-395, 696-700

Cell attachment sequence.

amino acids 428-431

Leucine zipper pattern.

amino acids 25-47

FIGURE 61

CTCCATTAAACCACCACCAGCTCCCCAAGCCACCCCTTCAGCC**ATG**AAGTTCCTGCTCCTGGT

CTTGGCAGCCCTCGGATTCCTGACCCAGGTGATCCCAGCCAGTGCAGGTGGGTCAAAATGTGT

GAGTAACACCCCAGGATACTGCAGGACATGTTGCCACTGGGGGGAGACAGCATTGTTCATGTG

CAACGCTTCCAGAAAATGCTGCATCAGCTACTCCTTCCTGCCGAAGCCTGACCTACCACAGCT

CATCGGTAACCACTGGCAATCAAGGAGAAGAAACACACAAAGGAAAGACAAGAAGCAACAAAC

GACCGTAACATCA**TAA**TAACCACTGCTATCGCCTCCACCAACTCAGAGAAATATCATTTCCAC

AGTTCCAATTCCTCCTACATTGCTGAGTACTAGCCAAGGCTCCTCTTTATGGGGCAGATATCT

ATAGCCAACCCCAAAACTTCTGTCTTCTATCATTCTGTCATTCATCTAGTAACTAATTTGGAG

TTTGTATCTATCTTACGAGAACAATCATCATGCAGATTCGTCCACAGGGGATCTGTCAGTTTG

GGTCCTCCAAATGAAAAATGTCAAGACAGAATTGGACATGCAAAGATTGACTGGGAGAACAC

ACCTCTGATGGACAAAGGTGAGACAGAGCAGCCACAGGCAGGGAGAGCCTTCAGACTGCAACG

CTGGCCTGATACGTGTCAAAGGAGAGAGGGATAGAGGAGGATTGAATAGAAGGAGACTAAGAC

TGCAGCTCTAAGAAAGTCTCAGCCAAACAGATGGGGAGGCCCAAAGCAAGGCTTGCCCCTCAG

AGGAGCTCACGCAGGGCAGGAATAGCCAGGTTCTCATATCCCAGGGGTTCAGACTTGGCTGAG

AACAGCCCCTGGAGAACATGGGGTGACTGCTACCATAGGTCTGGAAGTATGAGGCTGTCCACC

AACTATCCCCTTGAAGCAAGTTCTCTTGAAAGGAAATCTAAACAGTGCACCCCCATGGCTGCC

ACGGAGTATAAGGAGGGAGAGAAAGGAGCTGAAAGTCTAGGTTTGGCCAGCTAGGTAGACTGA

CTTGTGAGGTATTTATTTATTCATTTGAGTAACAAAGCAGACAGAATACATAGCCACCATTGG

TAGTACACCCCAAAAGCAAGGATGGCATGATGCTGGTGACTCAAACGTGCCTACTCATGGTGT

CAAATTGGCATAATCCTCTTGGGAAGCTGTGTGGAAATAAGCACAGAGAAGCAGAACTCTAAT

TGCTTAATCCACTAAACATTACTTCTGGGAATTGGCTCATCATAAATTATCCAAGAGAAAGCA

CAAAGTTATGGGCACAAAGGTTTTCCATATAATATTATTTAAAATGCTGAGAAAATGAAAAAA

TCTAAATGGTGAAATATATACTAATGCCATCTATAAATACAAACAAATAGAATGTTTATAGAA

TAATGGAACATAATAACATTATTCAAAATTGCATTTATGCTATAGTTGTCAAAATTGTCTCCT

TATATGATACAAAACTCATGAAAATTATGACTTTTTTGTTTGGTTGGAAAGCAGAATTATGCA

TAAATTTCCTCTTACAGTTCGATGCCCATTAGTTTTATATAACATTTATTTGACACGTACTGA

CTTCTATCTGAGAAGAACAAACCAAAACACTCAGGCCTAAATAATTAAAAACGGTCCTAAAAA

CTAGCAAACCAGATAAGAAAGATGTTAATGCCCATTCCCTAACTTATGTCTTAGACCAAAAT

TAATTCTAGATGGTTTTAAAATGACAGTGTAAAGTAAAGTATTAAAGATTGTGTGGTCAAA

TATTCAATTTAAGAGCAAGGAAATTCTTATAAATATAACAATAGAGGCAGAACTCATGTAAGA

ATAAATTGATTAGGTGGTATTAAATATTAAGTTCTTATGTATGTCAAAGATATCATTTTGAA

ATTCATCCATCTTATTGGGTATTGCAGGAGTTCATTCCTTTTTGTTTATAAATACTCTTCCGT

CATATGAATAGTATTCATTTGTATACTGGTTTGTTGATGGACATTTGGGTTGTTCCCAGTTTA

TGGCTATTACAAATAAAGCTTCTATGAACATTTATGTACA

FIGURE 62

MKFLLLVLAALGFLTQVIPASAGGSKCVSNTPGYCRTCCHWGETALFMCNASRKCCISYSFLP
KPDLPQLIGNHWQSRRRNTQRKDKKQQTTVTS

Important features of the protein:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 1-22

N-glycosylation site.

amino acids 50-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 79-82

N-myristoylation site.

amino acids 23-28

FIGURE 63

```
GCGGAGCGCCTGGGAGAGGAGAAGGAGCCGACCTGCCGAGATGGAGGCGACCGGCACCTGGGC
GCTGCTGCTGGCGCTGGCGCTGCTCCTGCTGCTGACGCTGGCGCTGTCCGGGACCAGGGCCCG
AGGCCACCTGCCCCCCGGGCCCACGCCGCTACCACTGCTGGGAAACCTCCTGCAGCTACGGCC
CGGGGCGCTGTATTCAGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGTGTTCACCATCTA
CCTGGGACCCTGGCGGCCTGTGGTGGTCCTGGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGG
AGGTCAGGCTGAGGAGTTCAGCGGCCGGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGG
CCATGGGGTTTTCTTCTCCAACGGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGC
TCTGCGGGACCTGGGCATGGGGAAGCGAGAAGGCGAGGAGCTGATCCAGGCGGAGGCCCGGTG
TCTGGTGGAGACATTCCAGGGGACAGAAGGACGCCCATTCGATCCCTCCCTGCTGCTGGCCCA
GGCCACCTCCAACGTAGTCTGCTCCCTCCTCTTTGGCCTCCGCTTCTCCTATGAGGATAAGGA
GTTCCAGGCCGTGGTCCGGGCAGCTGGTGGTACCCTGCTGGGAGTCAGCTCCCAGGGGGGTCA
GACCTACGAGATGTTCTCCTGGTTCCTGCGGCCCCTGCCAGGCCCCACAAGCAGCTCCTCCA
CCACGTCAGCACCTTGGCTGCCTTCACAGTCCGGCAGGTGCAGCAGCACCAGGGGAACCTGGA
TGCTTCGGGCCCCGCACGTGACCTTGTCGATGCCTTCCTGCTGAAGATGGCACAGGAGGAACA
AAACCCAGGCACAGAATTCACCAACAAGAACATGCTGATGACAGTCATTTATTTGCTGTTTGC
TGGGACGATGACGGTCAGCACCACGGTCGGCTATACCCTCCTGCTCCTGATGAAATACCCTCA
TGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGCACCAAGCCT
AGGGGACCGTACCCGCCTCCCTTACACCGACGCGGTTCTGCATGAGGCGCAGCGGCTGCTGGC
GCTGGTGCCCATGGGAATACCCGCACCCTCATGCGGACCACCCGCTTCCGAGGGTACACCCT
GCCCCAGGGCACGGAGGTCTTCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAA
GCACCAGAAGAGTTCAACCCAGACCGTTTCCTGGATGCAGATGGACGGTTCAGGAAGCATGA
GGCGTTCCTGCCCTTCTCCTTAGGGAAGCGTGTCTGCCTTGGAGAGGGCCTGGCAAAAGCGGA
GCTCTTCCTCTTCTTCACCACCATCCTACAAGCCTTCTCCCTGGAGAGCCCGTGCCCGCCGGA
CACCCTGAGCCTCAAGCCCACCGTCAGTGGCCTTTTCAACATTCCCCCAGCCTTCCAGCTGCA
AGTCCGTCCCACTGACCTTCACTCCACCACGCAGACCAGATGAAGGAAGGCAACTTGGAAGTG
GTGGGTGCCCAGGACGGTGCCTCCAGCCTCAACAGTGGGCATGGACAGGGTTAATGTCTCCAG
AGTGTACACTGCAGGCAGCCACATTTACACGCCTGCAGTTGTTTTCCGGAGTCTGTCCCACGG
CCCACACGCTCACTTGACTCATGCTGCTAAGATGCACAACCGCACACCCATACACAACTACAA
GGGCCACAAAGCAACTGCTGGGTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCAC
AAGCACATAGCCAGGTAACCCACCAACTCCCCTGGATCTGCAGCCCACACGTGGGAGTCTGGC
TGTCACCTTCACAAGCCACAGAAACGGCCACACATGTTCACAGCTCACACGCCCTCTCCATTC
ATCGAACTTCTCAGTGTCCCTGTCCCTGGTGCCTGGCACAGGGAACAGCATGCCCCCTCCGGG
GTCATGCCACCCAGAGACTGTCGCTGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACA
CCACTCTCCCAGCCTGTGACCACCGATGTCCACACACCCCCAACCACTTGTCCACACAGCTAC
CCACGTACAACATCGTCCTGGCTCCCCAGAGTATCTTCCACTGAGACACGCCGCCCCCACAG
AGGCACAGTCCCCAGCCACCTCTGCAACTGCAGCCCTCAGTCACCCCTTTTTAAGCACCCTGA
TTCTACCAAATGCAAACACATCTGGGTCTGCGATTATGCACAGAGACTTTGGACATACGAGGA
CCCTCAGACCGGAGGAACACCTGCCCAACCCCAACACGTGCTTATGTAACCACGTGGAAAGCG
GCCCCTGCTGCCCCTCCACACACACATACACACTCACTGATCTACAGCCCCTGTTCGGCGTCA
GAGTCCCCACTAGACCCAGTGGAAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACC
CTGTCAGGGAGTGAGCCGGATCTGACGTTCCTTGTGACTTAAGGGTCCGGCTTGGGAATTAAA
GTTTGTTTCTGGCCTTTAGCCTAAAAAAAAAAAAAAAAAAAA
```

FIGURE 64

MEATGTWALLLALALLLLLTLALSGTRARGHLPPGPTPLPLLGNLLQLRPGALYSGLMRLSKK
YGPVFTIYLGPWRPVVVLVGQEAVREALGGQAEEFSGRGTVAMLEGTFDGHGVFFSNGERWRQ
LRKFTMLALRDLGMGKREGEELIQAEARCLVETFQGTEGRPFDPSLLLAQATSNVVCSLLFGL
RFSYEDKEFQAVVRAAGGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHHVSTLAAFTVRQV
QQHQGNLDASGPARDLVDAFLLKMAQEEQNPGTEFTNKNMLMTVIYLLFAGTMTVSTTVGYTL
LLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDAVLHEAQRLLALVPMGIPRTLMRT
TRFRGYTLPQGTEVFPLLGSILHDPNIFKHPEEFNPDRFLDADGRFRKHEAFLPFSLGKRVCL
GEGLAKAELFLFFTTILQAFSLESPCPPDTLSLKPTVSGLFNIPPAFQLQVRPTDLHSTTQTR

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 294-313

Glycosaminoglycan attachment site.

amino acids 99-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 128-132

N-myristoylation sites.

amino acids 51-57, 109-115, 115-121, 188-194, 207-213, 257-263, 284-290, 339-345, 370-376, 444-450

Amidation sites.

amino acids 140-144, 435-439

Leucine zipper pattern.

amino acids 32-54, 39-61

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 433-443

FIGURE 65

CGGACGCGTGGGGCCGT**ATG**CGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCC

CCAGCCCCCTGCTCCTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGA

AGACCCGCCAGGTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTC

ATATACGGGCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGG

CAGTGGTAATGGTGGCCACCAACACCCCCACAGCACCCTGAGCATCAACTGGAGCCTCCTGC

TATCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCTG

CCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCAAAGC

CTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATCACTGATT

CATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGACCCTACCAGGA

CTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCCAGCCGACCAGCCC

AACCCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTGGCCCTGATTGGAGCCT

CTCCCCGGGGAAACCGTTCCCTGTTTGGGCTGGAGGTAGCCACATTGGGCCAGGGCCCTGACT

GCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATATGCACCGGCCGTCTTCCAGTTGG

ACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCACAGTGGCGACCAGTGGCTTACTCCC

AGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGCCAAGCTTCCCCTCTTCATCCTGCCTTAG

CATACTCTCTTCCCCAGTCACCCATTGTCCGAGCCTTCTTTGGGTCCCAGAATAACTTCTGTG

CCTTCAATCTGACGTTCGGGGCTTCCACAGGCCCTGGCTATTGGGACCAACACTACCTCAGCT

GGTCGATGCTCCTGGGTGTGGGCTTCCCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCA

TCATGGCAGTGGCCCTGGGTGCCCCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGC

TGCACCACAAGAAGTACTCAGAGTACCAGTCCATAAAT**TAA**GGCCCGCTCTCTGGAGGGAAGG

ACATTACTGAACCTGTCTTGCTGTGCCTCGAAACTCTGGAGGTTGGAGCATCAAGTTCCAGCC

GGCCCCTTCACTCCCCCATCTTGCTTTTCTGTGGAACCTCAGAGGCCAGCCTCGACTTCCTGG

AGACCCCCAGGTGGGGCTTCCTTCATACTTTGTTGGGGGACTTTGGAGGCGGGCAGGGGACAG

GGCTATTGATAAGGTCCCCTTGGTGTTGCCTTCTTGCATCTCCACACATTTCCCTTGGATGGG

ACTTGCAGGCCTAAATGAGAGGCATTCTGACTGGTTGGCTGCCCTGGAAGGCAAGAAAATAGA

TTTATTTTTTTTCACAGGGAAAAAAAAAAAA

FIGURE 66

MRGSVECTWGWGHCAPSPLLLWTLLLFAAPFGLLGEKTRQVSLEVIPNWLGPLQNLLHIRAVG

TNSTLHYVWSSLGPLAVVMVATNTPHSTLSINWSLLLSPEPDGGLMVLPKDSIQFSSALVFTR

LLEFDSTNVSDTAAKPLGRPYPPYSLADFSWNNITDSLDPATLSATFQGHPMNDPTRTFANGS

LAFRVQAFSRSSRPAQPPRLLHTADTCQLEVALIGASPRGNRSLFGLEVATLGQGPDCPSMQE

QHSIDDEYAPAVFQLDQLLWGSLPSGFAQWRPVAYSQKPGGRESALPCQASPLHPALAYSLPQ

SPIVRAFFGSQNNFCAFNLTFGASTGPGYWDQHYLSWSMLLGVGFPPVDGLSPLVLGIMAVAL

GAPGLMLLGGGLVLLLHHKKYSEYQSIN

N-glycosylation sites:

amino acids 65-69, 95-99, 134-138, 159-163, 187-191, 230-234, 333-337

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 397-401

Casein kinase II phosphorylation sites:

amino acids 151-155, 249-253, 255-259

N-myristoylation sites:

amino acids 3-9, 63-69, 235-241, 273-279, 292-298, 324-330

Leucine zipper pattern.

amino acids 371-393

FIGURE 67

CGGGACAGGCGCGTGAGGCCACAACACATGCGTGTATCTTGCTTGGGCTATCTTCCCTGCTCTGCCACGCCGGGT
CTGGAGAAGGGGTTTCAGCCCCAGGACATTTACTGAGAGTCGGCGAATATTGGGAGCCGCG**ATG**TTCCCCCTTCG
GGCCCTGTGGTTGGTCTGGGCGCTTCTAGGAGTGGCCGGATCATGCCCGGAGCCGTGCGCCTGCGTGGACAAGTA
CGCTCACCAGTTCGCGGACTGCGCTTACAAAGAGTTGCGTGAGGTGCCGGAAGGACTGCCTGCCAACGTGACGAC
GCTTAGTCTGTCCGCGAACAAGATCACTGTGCTGCGGCGCGGGGCCTTCGCCGACGTCACACAGGTCACGTCGCT
GTGGCTGGCGCACAATGAGGTGCGCACCGTGGAGCCAGGCGCACTGGCCGTGCTGAGTCAGCTCAAGAACCTCGA
TCTGAGCCACAACTTCATATCCAGCTTTCCGTGGAGCGACCTGCGCAACCTGAGCGCGCTGCAGCTGCTCAAAAT
GAACCACAACCGCCTGGGCTCTCTGCCCCGGGACGCACTCGGTGCGCTACCCGACCTGCGTTCCCTGCGCATCAA
CAACAACCGGCTGCGTACGCTGGCGCCTGGCACCTTCGACGCGCTTAGCGCGCTGTCACACTTGCAACTCTATCA
CAATCCCTTCCACTGCGGCTGCGGCCTTGTGTGGCTGCAGGCCTGGGCCGCGAGCACCCGGGTGTCCTTACCCGA
GCCCGACTCCATTGCTTGTGCCTCGCCTCCCGCGCTGCAGGGGGTGCCGGTGTACCGCCTGCCCGCCCTGCCCTG
TGCACCGCCCAGCGTGCATCTGAGTGCCGAGCCACCGCTTGAAGCACCCGGCACCCCACTGCGCGCAGGACTGGC
GTTCGTGTTACACTGCATCGCCGACGGCCACCCTACGCCTCGCCTGCAATGGCAACTTCAGATCCCCGGTGGCAC
CGTAGTCTTAGAGCCACCGGTTCTGAGCGGGGAGGACGACGGGGTTGGGGCGGAGGAAGGAGAGGGAGAAGGAGA
TGGGGATTTGCTGACGCAGACCCAAGCCCAAACGCCGACTCCAGCACCCGCTTGGCCGGCGCCCCCAGCCACACC
GCGCTTCCTGGCCCCTCGCAAATGGCTCCCTGTTGGTGCCCCTCCTGAGTGCCAAGGAGGCGGGCGTCTACACTTG
CCGTGCACACAATGAGCTGGGACTGGGCGCCAACTCTACGTCAATACGCGTGGCGGTGGCAGCAACCGGGCCCCCAAAACA
CGCGCCTGGCGCCGGGGGAGAACCCGACGGACAGGCCCCGACCTCTGAGCGCAAGTCCACAGCCAAGGGCCGGGG
CAACAGCGTCCTGCCTTCCAAACCCGAGGGCAAAATCAAAGGCCAAGGCCTGGCCAAGGTCAGCATTCTCGGGGA
GACCGAGACGGAGCCGGAGGAGGACACAAGTGAGGGAGAGGAGGCCGAAGACCAGATCCTCGCGGACCCGGCGGA
GGAGCAGCGCTGTGGCAACGGGGACCCCTCTCGGTACGTTTCTAACCACGCGTTCAACCAGAGCGCAGAGCTCAA
GCCGCACGTCTTCGAGCTGGGCGTCATCGCGCTGGATGTGGCGGAGCGCGAGGCGCGGGTGCAGCTGACTCCGCT
GGCTGCGCGCTGGGGCCCTGGGCCCGGCGGGCTGGCGGAGCCCCGCGACCCGGGCGGCGACCCCTGCGCCTACT
CTATCTGTGTCCAGCGGGGGGCGGCGCGGCAGTGCAGTGGTCCCGCGTAGAGGAAGGCGTCAACGCCTACTGGTT
CCGCGGCCTGCGGCCGGGTACCAACTACTCCGTGTGCCTGGCGCTGGCGGGCGAAGCCTGCCACGTGCAAGTGGT
GTTTTCCACCAAGAAGGAGCTCCCATCGCTGCTGGTCATAGTGGCAGTGAGCGTATTCCTCCTGGTGCTGGCCAC
AGTGCCCCTTCTGGGCGCCGCCTGCTGCCATCTGCTGGCTAAACACCCGGGCAAGCCCTACCGTCTGATCCTGCG
GCCTCAGGCCCCTGACCCTATGGAGAAGCGCATCGCCGCAGACTTCGACCCGCGTGCTTCGTACCTCGAGTCCGA
GAAAAGCTACCCGGCAGGCGGCGAGGCGGGCGGCGAGGAGCCAGAGGACGTGCAGGGGGAGGGCCTTGATGAAGA
CGCGGAGCAGGGAGACCCAAGTGGGGACCTGCAGAGAGAGGAGAGCCTGCGGCCTGCTCACTGGTGGAGTCCCA
GTCCAAGGCCAACCAAGAGGAGTTCGAGGCGGGCTCTGAGTACAGCGATCGGCTGCCCCTGGGCGCCGAGGCGGT
CAACATCGCCCAGGAGATTAATGGCAACTACAGGCAGACGGCAGGC**TGA**ACCTCCGCCCGTCCGGCCCGCCCATT
CCCGACCTCCACCTAGGGTGCCTGGGAGCAGCAGTCTAGGGCTGGCAGGACTTATGTCCCCCGTCCCCAACCTTC
ACCTACTCCTCCCCCTTACTACTCCCCAACCTTGACTACCAGGGACTTCTATTAGGGAGTGGGCCGATTTCACCA
GTCCCTGCTACCCACGGCTGCCATTCTCCCTGCGGGCTGAATCCCCTTCCCCGCCAAGCACAGTGTTTATCTTAC
CCCATGCAAGACTCCACCCGCAGACGGTGGGCGATATCTATGTCCCTCCATTCCCGTCGCGATTATCTGCGAAAT
CCACCCCGCAGCCCGCCCCACCGTGGGCTCTGGAGCCAGAGGAAACGAGCGAAGACTTTGGAAACCTCGCGGTAA
CGCGGTGGTTTCGGGGGCCAGCCAAGGCCAGTGGAGTGCTGTGGGGTCCCACCTCGACCCCTCCTCCTCCCTTTC
TTTCTTTCCTTTTTTTTTATTTTTTAATTTTTATTTATTTATTTATTTTTTGACGGAGTCTTGGTCTGTCGC
CAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCATCTTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTC
AGCCTGCCTAGTAGCTGGGACTACAGGCGCGCGCCACCACGACCAGCTAATTTCTTCTATTTTTAGTAGAGACGG
GGTTTCACCATGTTGGCCAGGATGGTCTGGATCTCTTGACCTCAGGTGATCCATCTGCCTCGGCCTCTCAAAGTG
CTGGGATTACAGGCGTGAGGCACCGCGCCCGGCCCCTCCTCCCTTTCAATCCCTACTCCCAGAAGCCGGGATTCG
TGGCAACCCCTAGTTTTTAGTTCCAAAGCCTCCTGCCGGCAGGGAACCAAATCCTTCTGTCCTCCCACCCCCACC
CCACTTCTGGCCAGTTGGAGTCCAGCCCGGTGCCTGGGGCGCCTTTCAGCTCCGCGCTCAGATTTTCCTGTTTTC
GTTGTTTTCAAAGACAGCGACATTTCGGGTCTGGTGCTAACACCCCCTTCCCAGCCTCTGGGAAAATCGAGTGTG
TGTGTCGGGGGGTAGGGAGGGAATGCGTTTTCTGTCGTCTCTCTCCTAACTTAAAGCGCCGCAGGACCGCGCGCC
CCTTGGCGGCTGAGCCTGTGGACTTGGTCGCGGGCCAATTTCGTTGTCCGTGTGTTGGGCTTTCCGGAGGTCTGT
GCGCCCAACAGCGCCGCTCCCGCGGCTCCACCCGACCCAGACCCTAGCTGGAAAGCGCCGGAGGCGGAGGAAGCT
GACTGTGGCCTCCCGGGCCGCGGCTCTCTGGAGGGCTCGCGCCCTAGTTCGCACAAAGCCTGCTCGTGACTGTGC
GACTGTGCGACGGGATCCGGATGGAGCCGAGCCCCTCCGTCCTCGCGTCTCGGTCCTCGCGTCGCCCCGCCCCAC
CCGCCCCTGCTTCGGCGGGAATCGTGTTTGCCCGGCGTGTAGTCCCTGACAAGCGTGCCCTGTAGGAGAAAAGTC
TGTGTCCTGTGAAGTGTGACCGTGTAGTGTAGGGGGCGGGCGGGGGGCGGATGGGCGGGGAGGGAGGGAAGGG
GAGGGGCGCGGCGCGGCGACTCGGGGCGGGGTTCTTTTTTCCATTTTGAAAGAAAGCGTCGGGGTTGGGGTGGGG
GGAGTTTCAGTCCTCGGGATCAGCCCTCTCCGCGAAGCGCAGCACAAGCGCGGGCCTGGGACGGAGTAGCCCCCC
GGAGCCCGTGCCCTTTTCTAAACGCGTCTGTATGCAGTCAATAAAACAATCGATTTGAAA

FIGURE 68

MFPLRALWLVWALLGVAGSCPEPCACVDKYAHQFADCAYKELREVPEGLPANVTTLSLSANKI
TVLRRGAFADVTQVTSLWLAHNEVRTVEPGALAVLSQLKNLDLSHNFISSFPWSDLRNLSALQ
LLKMNHNRLGSLPRDALGALPDLRSLRINNNRLRTLAPGTFDALSALSHLQLYHNPFHCGCGL
VWLQAWAASTRVSLPEPDSIACASPPALQGVPVYRLPALPCAPPSVHLSAEPPLEAPGTPLRA
GLAFVLHCIADGHPTPRLQWQLQIPGGTVVLEPPVLSGEDDGVGAEEGEGEGDGDLLTQTQAQ
TPTPAPAWPAPPATPRFLALANGSLLVPLLSAKEAGVYTCRAHNELGANSTSIRVAVAATGPP
KHAPGAGGEPDGQAPTSERKSTAKGRGNSVLPSKPEGKIKGQGLAKVSILGETETEPEEDTSE
GEEAEDQILADPAEEQRCGNGDPSRYVSNHAFNQSAELKPHVFELGVIALDVAEREARVQLTP
LAARWGPGPGGAGGAPRPGRRPLRLLYLCPAGGGAAVQWSRVEEGVNAYWFRGLRPGTNYSVC
LALAGEACHVQVVFSTKKELPSLLVIVAVSVFLLVLATVPLLGAACCHLLAKHPGKPYRLILR
PQAPDPMEKRIAADFDPRASYLESEKSYPAGGEAGGEEPEDVQGEGLDEDAEQGDPSGDLQRE
ESLAACSLVESQSKANQEEFEAGSEYSDRLPLGAEAVNIAQEINGNYRQTAG

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 587-610

N-glycosylation sites.

amino acids 52-55, 121-124, 337-340, 364-367, 474-477, 563-566

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 397-400

Casein kinase II phosphorylation sites.

amino acids 19-23, 202-205, 289-292, 246-249, 411-414, 431-434, 433-436, 440-443, 544-547, 583-586, 650-653, 700-703

N-myristoylation sites.

amino acids 15-20, 48-53, 165-170, 296-301, 351-356, 362-367, 390-395, 419-424, 514-519, 536-541, 557-562, 561-566, 610-615, 661-666, 716-721

Amidation site.

amino acids 522-525

Prokaryotic membrane lipoprotein lipid attachment sites.

amino acids 10-20, 603-613

FIGURE 69

```
GGCGGCGGGAGCAGCGAAGGGGGCGGCAGGGATCCTCCAGGCTGCCGGCTGGGAAGGCGTGGG
CGACCCGGTGTGTGGCGCGCCCAGAGCCCCGCGTTTCAGCCCTAGGGAAGGAAGCCAGTTGAG
GGAAGTTCTCCATGAATGTACGTCACAATGATGATGACCGACCAAATCCCTCTGGAACTGCCA
CCATTGCTGAACGGAGAGGTAGCCATGATGCCCCACTTGGTGAATGGAGATGCAGCTCAGCAT
GTTATTCTCGTTCAAGTTAATCCAGGTGAGACTTTCACAATAAGAGCAGAGGATGGAACACTT
CAGTGCATTCAAGGACCTGCTGAAGTTCCCATGATGTCACCCAATGGATCCATTCCTCCCATT
CATGTGCCTCCAGGTTATATCTCACAGGTGATTGAAGATAGTACTGGAGTCCGCCGGGTGGTG
GTCACACCCCAGTCTCCTGAGTGTTATCCCCCAAGCTACCCCTCAGCCATGTCTCCAACCCAT
CATCTCCCTCCCTATCTGACTCACCATCCACATTTTATTCATAACTCACACACGGCTTACTAC
CCACCTGTTACCGGACCTGGAGATATGCCGCCTCAGTTTTTTCCCCAGCATCATCTTCCCCAC
ACAATATATGGTGAGCAAGAAATTATACCATTTTATGGAATGTCAAGCTACATCACCCGAGAA
GACCAGTACAGCAAGCCTCCGCACAAAAAACTGAAAGACCGCCAGATCGATCGCCAGAACCGC
CTCAACAGCCCTCCTTCTTCTATCTACAAAAGCAGCTGCACAACAGTATACAATGGCTATGGG
AAGGGCCATAGTGGTGGAAGTGGCGGAGGCGGCAGCGGTAGTGGTCCCGGAATTAAGAAAACA
GAGCGACGAGCAAGAAGCAGCCCAAAGTCGAATGATTCAGACTTGCAAGAATATGAGTTGGAA
GTAAAGAGGGTGCAAGACATTCTTTCGGGAATAGAGAAACCACAGGTTTCTAATATTCAGGCA
AGAGCAGTTGTGTTGTCCTGGGCTCCCCCTGTTGGACTTTCCTGTGGACCCCACAGTGGTCTT
TCCTTCCCCTACAGTTACGAGGTGGCCTTATCAGACAAAGGACGAGATGGAAAATACAAGATA
ATTTACAGTGGAGAAGAATTAGAATGTAACCTGAAAGATCTTAGACCAGCAACAGATTATCAT
GTGAGGGTGTATGCCATGTACAATTCCGTAAAGGGATCCTGCTCCGAGCCTGTTAGCTTCACC
ACCCACAGCTGTGCACCCGAGTGTCCTTTCCCCCCTAAGCTGGCACATAGGAGCAAAAGTTCA
CTAACCCTGCAGTGGAAGGCACCAATTGACAACGGTTCAAAAATCACCAACTACCTTTTAGAG
TGGGATGAGGGAAAAGAAATAGTGGTTTCAGACAGTGCTTCTTCGGGAGCCAGAAGCACTGC
AAGTTGACAAAGCTTTGTCCGGCAATGGGGTACACATTCAGGCTGGCCGCTCGAAACGACATT
GGCACCAGTGGTTATAGCCAAGAGGTGGTGTGCTACACATTAGGAAATATCCCTCAGATGCCT
TCTGCACCAAGGCTGGTTCGAGCTGGCATCACATGGGTCACGTTGCAGTGGAGTAAGCCAGAA
GGCTGTTCACCCGAGGAAGTGATCACCTACACCTTGGAAATTCAGGAGGATGAAAATGATAAC
CTTTTCCACCCAAAATACACTGGAGAGGATTTAACCTGTACTGTGAAAAATCTCAAAGAAGC
ACACAGTATAAATTCAGGCTGACTGCTTCTAATACGGAAGGAAAAGCTGTCCAAGCGAAGTT
CTTGTTTGTACGACGAGTCCTGACAGGCCTGGACCTCCTACCAGACCGCTTGTCAAAGGCCCA
GTTACATCTCATGGCTTTAGTGTCAAATGGGATCCCCCTAAGGACAATGGTGGTTCAGAAATC
CTCAAGTACTTGCTAGAGATTACTGATGGAAATTCTGAAGGTGAAGTTTTTGGCAATTGTTTT
ATTCAAATCCAATAGCAAGCTCTGTTTTCTAATATAGTAAATGTCTTTATAGTAATAGTGAGT
AATCATTAATTCTAAAGATAGAATTATTATTACAATAAACAAACTTTAGTCACATATTGGCAG
TTTTTCTATTTCAAACACAGCACCAGAGATCAGAGTCTACTTGAAACTTACATTTGTGTTATT
TAACAATTTTTCTGTATCTTTTTCATTGGTGTTTTGTTTTGTTTATCTTTTGTTTTTGTTTCT
TTGGTTTGGTTTGTTTTTGTTTTGTTTTTTGAGATACGATCTCTGTCACACAGGCTGGAGGGC
AGTGGCACAGACATGGCCCATTGCAGTCTCAGACTCCTGGGCTTAAGTGACTCTTCTGCCACA
GAAGATGAGGAAGAATACATTTTTCATAGTGATGGGGTCTCACTATGTTATCTAGGCTGGTCT
CAAACTCCTGGCCTCAAGCAACCCTCCACCTTGGCCTCCCAAAGTGCTGGGACTATAGACATG
AATCACCACACTCAGCTTCCATGTCTTTTTATGAACTAGGGTTCCTAATTAATCAGATAAATT
TGGTATTTTCATCTCCTAACTTGCCATATGTTTTCTGGAATTCTTATAAGCAGCCGAGAGTG
GTGGCTCACGCTGTAGTCCCAGCACTTTGGGAGGCTGAGGTGGGTGGTCAGGAGATCAAGACC
ATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGTGTGGTG
GCAGGCACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGAAGAATTGCTTGAACCCAGCAG
GCGGAGGTTGCAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGTGACAGAGTGAGACTC
TGTCTCAAAAAAAAAAAAA
```

FIGURE 70

MMMTDQIPLELPPLLNGEVAMMPHLVNGDAAQHVILVQVNPGETFTIRAEDGTLQCIQGPAEV
PMMSPNGSIPPIHVPPGYISQVIEDSTGVRRVVVTPQSPECYPPSYPSAMSPTHHLPPYLTHH
PHFIHNSHTAYYPPVTGPGDMPPQFFPQHHLPHTIYGEQEIIPFYGMSSYITREDQYSKPPHK
KLKDRQIDRQNRLNSPPSSIYKSSCTTVYNGYGKGHSGGSGGGGSGSGPGIKKTERRARSSPK
SNDSDLQEYELEVKRVQDILSGIEKPQVSNIQARAVVLSWAPPVGLSCGPHSGLSFPYSYEVA
LSDKGRDGKYKIIYSGEELECNLKDLRPATDYHVRVYAMYNSVKGSCSEPVSFTTHSCAPECP
FPPKLAHRSKSSLTLQWKAPIDNGSKITNYLLEWDEGKRNSGFRQCFFGSQKHCKLTKLCPAM
GYTFRLAARNDIGTSGYSQEVVCYTLGNIPQMPSAPRLVRAGITWVTLQWSKPEGCSPEEVIT
YTLEIQEDENDNLFHPKYTGEDLTCTVKNLKRSTQYKFRLTASNTEGKSCPSEVLVCTTSPDR
PGPPTRPLVKGPVTSHGFSVKWDPPKDNGGSEILKYLLEITDGNSEGEVFGNCFIQIQ

Important features of the protein:

N-glycosylation sites.

amino acids 69-73, 254-258, 401-405

Glycosaminoglycan attachment sites.

amino acids 229-233, 234-238, 236-240

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 416-420, 535-539

Tyrosine kinase phosphorylation site.

amino acids 319-326

N-myristoylation sites.

amino acids 52-58, 227-233, 228-234, 230-236, 231-237, 232-238, 235-241, 239-245, 402-408, 610-616

Amidation site.

amino acids 414-418

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 290-301

ATP/GTP-binding site motif A (P-loop).

amino acids 546-554

CUB domain proteins profile.

amino acids 294-301

FIGURE 71

```
AAGTCATTCAGTGGATGTGATCTTGGCTCACAGGGGACGATGTCAAGCTCTTCCTGGCTCCTTCTCAGCCTTGTT
GCTGTAACTGCTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATTTTTGGACAAGTTTAACCACGAAGCCGAA
GACCTGTTCTATCAAAGTTCACTTGCTTCTTGGAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATG
AATAATGCTGGGGACAAATGGTCTGCCTTTTTAAAGGAACAGTCCACACTTGCCCAAATGTATCCACTACAAGAA
ATTCAGAATCTCACAGTCAAGCTTCAGCTGCAGGCTCTTCAGCAAAATGGGTCTTCAGTGCTCTCAGAAGACAAG
AGCAAACGGTTGAACACAATTCTAAATACAATGAGCACCATCTACAGTACTGGAAAAGTTTGTAACCCAGATAAT
CCACAAGAATGCTTATTACTTGAACCAGGTTTGAATGAAATAATGGCAAACAGTTTAGACTACAATGAGAGGCTC
TGGGCTTGGGAAAGCTGGAGATCTGAGGTCGGCAAGCAGCTGAGGCCATTATATGAAGAGTATGTGGTCTTGAAA
AATGAGATGGCAAGAGCAAATCATTATGAGGACTATGGGGATTATTGGAGAGGAGACTATGAAGTAAATGGGGTA
GATGGCTATGACTACAGCCGCGGCCAGTTGATTGAAGATGTGGAACATACCTTTGAAGAGATTAAACCATTATAT
GAACATCTTCATGCCTATGTGAGGGCAAAGTTGATGAATGCCTATCCTTCCTATATCAGTCCAATTGGATGCCTC
CCTGCTCATTTGCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACTCTTTGACAGTTCCCTTTGGACAG
AAACCAAACATAGATGTTACTGATGCAATGGTGGACCAGGCCTGGGATGCACAGAGAATATTCAAGGAGGCCGAG
AAGTTCTTTGTATCTGTTGGTCTTCCTAATATGACTCAAGGATTCTGGGAAAATTCCATGCTAACGGACCCAGGA
AATGTTCAGAAAGCAGTCTGCCATCCCACAGCTTGGGACCTGGGGAAGGGCGACTTCAGGATCCTTATGTGCACA
AAGGTGACAATGGACGACTTCCTGACAGCTCATCATGAGATGGGGCATATCCAGTATGATATGGCATATGCTGCA
CAACCTTTTCTGCTAAGAAATGGAGCTAATGAAGGATTCCATGAAGCTGTTGGGGAAATCATGTCACTTTCTGCA
GCCACACCTAAGCATTTAAAATCCATTGGTCTTCTGTCACCCGATTTTCAAGAAGACAATGAAACAGAAATAAAC
TTCCTGCTCAAACAAGCACTCACGATTGTTGGGACTCTGCCATTTACTTACATGTTAGAGAAGTGGAGGTGGATG
GTCTTTAAAGGGGAAATTCCCAAAGACCAGTGGATGAAAAAGTGGTGGGAGATGAAGCGAGAGATAGTTGGGGTG
GTGGAACCTGTGCCCCATGATGAAACATACTGTGACCCCGCATCTCTGTTCCATGTTTCTGATGATTACTCATTC
ATTCGATATTACACAAGGACCCTTTACCAATTCCAGTTTCAAGAAGCACTTTGTCAAGCAGCTAAACATGAAGGC
CCTCTGCACAAATGTGACATCTCAAACTCTACAGAAGCTGGACAGAAACTGTTGTAAGAAATACCTCAAATGTT
GAACCTCTCCTAGTATTCAGTATTACTCATTTCCATGCCTAGGTTTGTATTTGATTTCTTTGTTCTAAAAGAAA
ATTTTATGGCCTCAAAATGTCCTCATTTACAAACCAAACATTTAATTTGTGGTCAGACAGGAACCTAGACCATAC
AACAATTGGGTGGGCCACCTCTTTTCTCCCTATCATAACTACAGCCCTCTCTTCCTGGTAATTGGAAGGAAAGAG
CGGTTTAGGGTGGAATATATCTGTTAATATGCATTCTTTTCTTATCTGCCAGAAGCAAATTTAGCCAAGTCAAAG
AGAAGAAACCATAGATCATAGATGTAAATATATGTACATCTGGAACCCCTCAAAAGGCCCTGAACCCCCTTTTTT
TGTGTAGCAATATGCTGAGGCTTGGAAAATCAGAACCCTGGACCCTAGCATTGGAAAATGTTGTAGGAGCAAGAA
CATGAATGTAAGGCCACTGCTCAACTACTTTGAGCCCTTATTTACCTGGCTGAAAGACCAGAACAAGAATTCTTT
TGTGGATGGAGTACCGACTGGAGTCCATATGCAGACCCAAAGCATCAAAGTGAGGATAAGCCTAAAATCAGCTC
TTGGAGATAAAGCATATGAATGGAACGACAATGAAATGTACCTGTTCCGATCATCTGTTGCATATGCTATGAGGC
AGTACTTTTTAAAAGTAAAAAATCAGATGATTCTTTTTGGGGAGGAGGATGTGCGAGTGGCTAATTTGAAACCAA
GAATCTCCTTTAATTTCTTTGTCACTGCACCTAAAAATGTGTCTGATATCATTCCTAGAACTGAAGTTGAAAAGG
CCATCAGGATGTCCCGGAGCCGTATCAATGATGCTTTCCGTCTGAATGACAACAGCCTAGAGTTTCTGGGGATAC
AGCCAACACTTGGACCTCCTAACCAGCCCCCTGTTTCCATATGGCTGATTGTTTTTGGAGTTGTGATGGGAGTGA
TAGTGGTTGGCATTGTCATCCTGATCTTCACTGGGATCAGAGATCGGAAGAAGAAAAATAAAGCAAGAAGTGGAG
AAAATCCTTATGCCTCCATCGATATTAGCAAAGGAGAAAATAATCCAGGATTCCAAAACACTGATGATGTTCAGA
CCTCCTTTTAGAAAAATCTATGTTTTTCCTCTTGAGGTGATTTTGTTGTATGTAAATGTTAATTTCATGGTATAG
AAAATATAAGATGATAAAGATATCATTAAATGTCAAAACTATGACTCTGTTCAGAAAAAAAATTGTCCAAAGACA
ACATGGCCAAGGAGAGAGCATCTTCATTGACATTGCTTTCAGTATTTATTTCTGTCTCTGGATTTGACTTCTGTT
CTGTTTCTTAATAAGGATTTTGTATTAGAGTATATTAGGGAAAGTGTGTATTTGGTCTCACAGGCTGTTCAGGGA
TAATCTAAATGTAAATGTCTGTTGAATTTCTGAAGTTGAAAACAAGGATATATCATTGGAGCAAGTGTTGGATCT
TGTATGGAATATGGATGGATCACTTGTAAGGACAGTGCCTGGGAACTGGTGTAGCTGCAAGGATTGAGAATGGCA
TGCATTAGCTCACTTTCATTTAATCCATTGTCAAGGATGACATGCTTTCTTCACAGTAACTCAGTTCAAGTACTA
TGGTGATTTGCCTACAGTGATGTTTGGAATCGATCATGCTTTCTTCAAGGTGACAGGTCTAAAGAGAGAAGAATC
CAGGGAACAGGTAGAGGACATTGCTTTTTCACTTCCAAGGTGCTTGATCAACATCTCCCTGACAACACAAAACTA
GAGCCAGGGGCCTCCGTGAACTCCCCAGAGCATGCCTGATAGAAACTCATTTCTACTGTTCTCTAACTGTGGAGT
GAATGGAAATTCCAACTGTATGTTCACCCTCTGAAGTGGGTACCCAGTCTCTTAAATCTTTTGTATTTGCTCACA
GTGTTTGAGCAGTGCTGAGCACAAAGCAGACACTCAATAAATGCTAGATTTACAAAA
```

FIGURE 72

MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMN

NAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTILNTMSTI

YSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLYEEYVVLKNE

MARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAY

PSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFF

VSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTAHHEMGH

IQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLK

QALTIVGTLPFTYMLEKWRWMVFKGEIPKDQWMKKWWEMKREIVGVVEPVPHDETYCDPASLF

HVSDDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLL

Important features of the protein:
Signal peptide:
amino acids 1-17

N-glycosylation sites.
amino acids 53-57, 90-94, 103-107, 322-326, 432-438, 546-550

N-myristoylation sites.
amino acids 260-266, 286-292, 395-401

Cell attachment sequence.
amino acids 204-207

Neutral zinc metallopeptidases, zinc-binding region signature.
amino acids 371-381

FIGURE 73

CCCACGCGTCCGAGCGGGGTGGACAAGTGGCGTGTGTGCTGCGACCCCGAGGGAAG**ATG**AACG

GGACGCGGAACTGGTGTACCCTGGTGGACGTGCACCCAGAGGACCAGGCGGCGGCGGGCAGGA

AGACCTATGCCATGGTGTCCAGCCACTCAGCTGGTCATTCTCTGGCTTCAGAACTGGTGGAGT

CCCATGATGGACATGAGGAGATCATTAAGGTGTACTTGAAGGGGAGGTCTGGAGACAAGATGA

TTCACGAGAAGAATATTAACCAGCTGAAGAGTGAGGTCCAGTACATCCAGGAGGCCAGGAACT

GCCTACAGAAGCTCCGGGAGGATATAAGTAGCAAGCTTGACAGGAACCTAGGAGATTCTCTCC

ATCGACAGGAGATACAGGTGGTGCTAGAAAAGCCAAATGGCTTTAGTCAGAGTCCCACAGCCC

TGTACAGCAGCCCACCTGAGGTGGACACCTGTATAAATGAGGATGTTGAGAGCTTGAGGAAGA

CGGTGCAGGACTTGCTGGCCAAGCTTCAGGAGGCCAAGCGGCAACACCAGTCAGACTGTGTGG

CTTTTGAGGTCACACTCAGCCGGTACCAGAGGGAAGCAGAACAAAGTAATGTGGCCCTTCAGA

GAGAGGAGGACAGATGTCCAGAG**TGA**TTGGAGAATGTCCTGGGGGAATGAAGTTCCTTCCACA

AACACAGCTCAGTTCTTAGCAACAAACTGTTTGTTTTTCTACTTGCTCCATCTGCAGCCTACG

CTGCCCTGGCCTCCTGCAGACAGATAGTGGGGTTACCTGGCAAGGCCTGGTGAGAGCCAGTGA

ACCTAAGCTTTGACTGGGTGGCCTTGTCTTTCTGGGGAGGAGGGAATGTACATTCAGGGAGTA

GCCTTTTGCGGAAAAATTCTCTAGGGCTACAGACAGTCATGTGTGACTTCTCTCTGCTGTGAA

AACTCCCAGAGTCTCTTTAGGGATTTTCCCTAAGGTGTACCACCAGGCACACCTCAGTCTTCT

TGACCCAGAGCCTGAAAACTGTTTTCACTGGGTTCCACCAGTCCCAGCAAAATCCTCTTTGTA

TTTATTTTGCTAAGTTATTGGTGGTTTTGCTTACATCTCATGATTGATATAATACCAAAGTTC

TATAGCCTTCTCTTGCAGTATTTGGATTTGCTTGAAACCGGGAAAACTGTTCCCATTAGGCTT

GTTAATGTCAGAGTGACACTATTATGAATCTTTCTCTCCCTTTCCTCTGCCTGTTTCTTCTCT

CTTTCTCCTTCAAACTTGCTCTGCAGCTAAGGAAGGTGAGTCTACTTTCCCTGAGGCTTTGGG

GTCAGAGTATATGTTGTTTGGAGAAAGAGGGCAATCAGGACTCTTCTGGGACCCAGATGAGTT

CTTCACTAGCCCTTCTGAACCCCTTGCTCCATAATTGGTCTTTTATCCTGGCTCTGAATGACC

CTGCAGGTCATCATGGTTTTCTTTTTTTATTGTTTTTTTTTTTTCTGAGACAGAGTCTCACT

CTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCAGCTCACTGCAACCTCTGCCTCCCGGAT

TTAAGCGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGTGCCACCACGCCTG

GCTGATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATACTGGCTAGGCTGGTCTCGAATT

CCTGACCTCAGGTGATCCACCCACCTCGGCTTCCCAAAGTGCTAGGATTATAGGCTTGAGCTA

CTGCGCCCGGCCCATGGTGTTTTTCTTTAGGGCTCTTCCTACAGCCTTGAGAAGTAGATAGGC

ATCAGAGTATGGTACTATAGGAATCAGAAAAATTCAAAACAAATGTGGATTAAGTGTTTAGGC

TCTATGTGGCTCACGCAGCCAGAATCCTTAAGTCTGTGTGTTTCTGTGTCTCAAGACTGGGCT

CACATTCTGGCTTTGTCCATAACAATGCTCTGGGATTTCAGGGAGTTCCCTCATTTGTAAAAT

GAGGGGGTCAGAGCAGGTGATATCCATGTTTCTTCCCTTTCTGATATTGTTGTCTGTGGCATA

TTCTTTGTATGGCGAATTTAATAAATTATATTAATGTGTCA

FIGURE 74

MNGTRNWCTLVDVHPEDQAAAGRKTYAMVSSHSAGHSLASELVESHDGHEEIIKVYLKGRSGD
KMIHEKNINQLKSEVQYIQEARNCLQKLREDISSKLDRNLGDSLHRQEIQVVLEKPNGFSQSP
TALYSSPPEVDTCINEDVESLRKTVQDLLAKLQEAKRQHQSDCVAFEVTLSRYQREAEQSNVA
LQREEDRCPE

Important features of the protein:

Signal peptide:

amino acids 1-39

N-glycosylation site.

amino acids 2-6

Amidation site.

amino acids 21-25

FIGURE 75

```
GCTTGCACACATGGCTCCGGAGGCTCCGGTTGCCCATCCGAGCCCCTGCCAGGCTCTAACGTTCCCAACTGACAA
CACCAGTAACTAAATATAGGAGCAGATGGTGGGGACGGGCTGTCGCAGCGGCTCCTTTGCAGAGGTCTCCGGACT
GCAGATAAGGCTCAGGCCCTTTTGTGAGAAGCAGACCAGCCTGGGGGCTGGCGGCAGGACACCTGTGTCTGCATG
CTGAAGAAGATGGGTGAGGCCGTGGCCAGAGTAGCAAGGAAGGTCAACGAGACGGTGGAGAGCGGCTCTGACACT
CTGGACCTGGCCGAGTGCAAGCTGGTCTCCTTTCCCATTGGCATCTACAAGGTCCTGCGGAATGTCTCTGGCCAG
ATCCACCTCATCACCCTGGCTAACAACGAGCTTAAGTCCCTCACCAGCAAGTTCATGACCACATTCAGTCAGCTC
CGAGAGCTCCACCTGGAGGGGAACTTCCTACACCGCCTCCCCAGCGAGGTCAGTGCCCTGCAGCACCTCAAGGCC
ATTGACCTGTCCCGGAACCAGTTCCAGGACTTCCCTGAGCAGCTTACCGCCCTGCCGGCGCTGGAGACCATCAAC
CTGGAGGAGAACGAGATCGTAGATGTGCCCGTGGAGAAGCTGGCCGCCATGCCAGCCTTGCGCAGCATCAACCTC
CGCTTCAACCCACTCAACGCCGAGGTGCGCGTGATCGCCCCGCCGCTCATCAAGTTTGACATGCTCATGTCTCCG
GAAGGCGCAAGAGCCCCCCTACCTTAGGCCACCCTCCTCATGCCCACCCAGCAAGGGACAGAGGCCACAGGCCTG
GAACCCTGGAAGGGAGGGAGGCCCATGGGAGGCCAAGCCTGGGGGCTGGGGGCGGGTGGGCCGAGCAGCACGTGG
TGGGTGGGGTGCAGCTGGTCTGGATAGATAGCTTACAGCAGTAGTGGGCTCTGGAATGCCCAAGGGAAGAGGCAA
GGTGGGGCCTGCAGCCTGGACTCGGCACTCACAGCTGCTGTGCAAACTCAGGCAGATCTCCTGCCCTCTCTGAGC
CTTGTCACTTGAAAAAAACAGGACCCTTTCCCTCCTTTGGGCTCCCTGGAGGTTTTTAAGCAGTACGTGCCTCCA
AGTTACCTCCAGATCAGCAGGCACAGGTGGGCATTGCCAGGTATTTTCTGAGCCCCTGCGGGTTTGAGGCCTTGT
TTTTAGTGCTGAGAGCCAGTTGCTGCCCTGAGAAGAGAAGACAACCTCCATCTATTTATTGCTTCCTGAGAACTG
ACCTGGATGCGGCCCTCTGCAGGGCCCAGTCTTCAGTCCTGTGGTCCCTGGACTGGTGGGAACCTGAACTAGGAG
TCCTGGGAGAGCTGTGGTGGGAATATGGGCTGGCACTGCTGCAGGGCAAGAACATTCATGTAGGAGCCCGAGGAC
CANCANGCTGGGAATGGGGAGCAAGTCACGTCAGCTCTGTCATTCCCCACAGTTAACAAATTGGCGGGGTGGGAA
GTCCTGAGTGCTCCGTCCCTCTAGCATCACTCCTGAGCTGCGGGAGAGGTGGCCCAGAGAACAGCAGAGTCAGTT
ACACCTGCAGCTCTTGTCTAAAGTGATTAGATGGCCACCCTCACCACTGTCCAGTCCAGCAGCAGCCTGGCTGCC
TTGTCATGGCCTCCTGGGGGCAGAAGGCGATGTGGACCACGGGATTTGTAGCCAGCCAGCTCCCAGGCCAACGCC
CAAAGCCCTGATGACCTGGTTCTTCTGAGGCCCTCAACCTGGCATCTTAGGGTATGGTCAGGCAACAGGGTGACC
AGCTGTCCTGGTTTCCCAGGACATGGAACTTTCAATGCTAAAACTGGGACATTACCCAGCAAGTGGGGATGGTTG
GTCCCCTACCAGGAGAGGGCCTGGGGCTCTTGCTTCCCGAGAACGCCTGTGGCTTGAAGAACCTTGACTGCTTGG
TCCTCAGGTATCTACCTCCCACCTTCTCCTCATCTGTGGAGCAAGCCAACTCAGTGCCCCAGACCCCACCTGATC
TGCATCTTTGTTTGCTCCAGAGACACCTGAGGCCCCAGAGCTTGAGGCAAAGCCAGGCCGTCCAAATCCTGTGTG
CCGTGGACGAGTGGCCACTTTACTACTCCTAAGGCTAAGATGTTGAGAGCTCAGACCACTGCTCAGAGCAGTAAT
CCCTGCTCAGAATGCTCCCAGTTCCCTCGTCCCTGCCCAGGTCTCTTGTCTCTTGGGAAGGAACTGATAGGTCGG
GCCATTGTTGGGCCATCACTGAGCGCTCAGTATCTCAAGAGACTCTGTTCATTCTGCTCGTATCCCAAGGCCTGG
TTGGTCAAACTCTGGGCAAAGGGTTTTCAGGATGAGGAGGTCAAGACAGGATGTCCAGAGCTACCGAGTTCATCT
GTGGGTGTTGGGGGCAAGTGGGGGCTGAAGTCCTGTGCAGGCTGCGCTGGCCCCACCTGCCTTGTGCCCTGGAGT
GGGGTTTCTCCTTGTTGAAGAAGAGGCATCCTTCTCTGATGTGCACAAACACAATGTATGACCAGAGCCTTGCAA
CTCAAAGTGTGGTCTGTGGACCAGCAGCGGCAGTGACACCTGGGAGCTTGTTAGGAATGCAGAGTCTAGGCCTCA
CCCTATACCTCCCGACTCAGACCCTGCATTTTAGCAAGACCCCCAGCTGATTCCTATAAGCACTTTAGAGTTTGA
GAAGCAAGGACCTAGGCTGGGGATGTCCTCCGAGCAGAGGGTGAAGTTTCTCTCAGTTCTCTCCCTGCCACTTCC
AGGGATCTGAGCCTGTGTTCAGCCTCCTCCCTAACCCACCCTGGGAGACACTTGGCCTGTTAGATTGTTCCAGAG
TCTGCATGGCACTCCTGAAGAAGGGAGTGTGACCTGCAGTCACCAGGAGATGAGGGTTAGGTGTGCCCAGCCCTC
CAGACCCGGCCTTTCTGGTTAACCCCTGCATGCCAAGCTGCCTGCTGCCCCAGGTCCTCACCTCAGGCCTTTGAA
GGGGCAGCTTCTGGAAGTTGTTTTCTCCTCTGCTTGGAGAGTTTGCCCTTGTCTGTCTTGGAAAGTGTGGGCAGC
CACAGATGCCCCCAAATCAGAGCTCACAGTGAGTGAGCCCCTAAGCTTCAGTCTGCAATAAAGAATGCATTGGTT
TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 76

MLKKMGEAVARVARKVNETVESGSDTLDLAECKLVSFPIGIYKVLRNVSGQIHLITLANNELK
SLTSKFMTTFSQLRELHLEGNFLHRLPSEVSALQHLKAIDLSRNQFQDFPEQLTALPALETIN
LEENEIVDVPVEKLAAMPALRSINLRFNPLNAEVRVIAPPLIKFDMLMSPEGARAPLP

Important features of the protein:

N-glycosylation sites.

amino acids 17-21, 47-51

FIGURE 77

```
CACCAACAAGCAATCGTTCATGAGAAAGCCGTGCACCCGCTGCAGTTGGGCCATGTGGTCCGCATCGTATTCCAC
TAGGTCCCCATTGTACACCAAGTACTGTCCCGGCGTCTCCAGCAGATGCCTGCAGCCTTCCACCTTCTCAAGCAG
GGTGGTGTGAGTGCGCTGCTTTCCTTCCTCGCCTGGACCGGAGCCGTCGCGGGAGGCACCCCCGGGGGTGGAGAA
AAAGCCGGCCTGGCCTCGGAGGTGGTCTCGGCCCCCCGCCCCACCGACTCCCTCCTCCCCTCCAGAGGCGGCGGC
GGCTCCGGCGGCAGCAGCGGCAGGCAGCAACGTAAGCGGGATGCTCTCCAGGCTGCTTTTCTGCTCGGTCAGCAA
ATGGCTGAGCTGGTACATCTCGCTCTCCAGGTAGGAGATCTCGCGGGCCGTCTCTATGAACTGCCGGTAGTTCTG
GTAGACGTTGCGCTTCAGGTTCTGCGCCGTCTCCTCCGCCAGCGCCTGGATGCGCTGCCGGTGCTCCTGGAGGTC
CCGGTCCCCATCCGACTGCTGCGAGAGCTGCTTCACGTACAGCCGCGCCTCAAAACCCCCTGACTCCAGCTGCCG
ACGCAGGCGGCTCGCCCCACTGTCCGACATCGCCATCGCCATTTCTCTCCGGGTCTCACGCACTCACTGTCACTA
TCGGCGCCGCAGCCGCCGCGGCTGTCTAGACCCACCCAAGGCCAACCGAGCTCCTGGGCTGAGGAAGCAGGAATG
GGAACGAGACGAGTACGCCTGCGCCGGGTCTGAGCGTCAGACACTGCGCCTGCGCAAGTGGGCCGAGCGCAGACA
TTGCGCCTGCGCAGCAATGCCATCGGTTAAAGCGCATGCGCAAGATGAGCTATTGCGGAAGTGAGGGGAGGGAGA
GGCCGAGAGAAATTTCGGTACTGCGCATGAACCGAGCGTGACGTTGAGGTTTGAAATAACCGGCAAAGAGTAAAG
GCTGAAACTAGCTTCCTGAAAGCTTCGTAGGGCCCGAGCCCTGTGAGCCCAGGTTCTGCGCCCACTAGGAGGTGT
CATGCTGACTGCTTTTTTTAAAGCCCTAGAATCCTTGGCTTCGGCGTTTGGGGTAAGCTCCGTTCTCGTTCTCAA
GCGCGTTTCCGCGAACTCTCGCGGGATTGACGGGCCGTCTCGAGAGCCGGCATCTCCTAGGAGCTAGTCCTGGTC
CTCGGCTAGGCGGCTTGGGGTCGCGGCGTAACTGGGGAGCCAGCCTGACGCCGGCGGACCCCGCCTGTGATCCTG
GCAACG**ATG**GATGATGACTTGATGTTGGCACTGCGGCTTCAGGAGGAGTGGAACTTGCAGGAGGCGGAGCGCGAT
CATGCCCAGGAGTCCCTGTCGCTAGTGGACGCGTCGTGGGAGTTGGTGGACCCCACACCGGACTTGCAGGCACTG
TTTGTTCAGTTTAACGACCAATTCTTCTGGGGCCAGCTGGAGGCCGTCGAGGTGAAGTGGAGCGTGCGAATGACC
CTGTGTGCTGGGATATGCAGCTATGAAGGGAAGGGTGGAATGTGTTCCATCCGTCTCAGCGAACCCCTTTTGAAG
TTGAGGCCAAGAAAGGATCTTGTAGAGACCCTCCTGCATGAAATGATACATGCCTATTTATTTGTCACTAATAAC
GACAAAGACCGAGAAGGGCATGGTCCAGAATTTTGTAAACATATGCATCGCATCAACAGCCTGACTGGAGCCAAT
ATAACGGTATACCATACTTTTCACGATGAGGTGGATGAGTATCGGCGACACTGGTGGCGCTGCAATGGGCCGTGC
CAGCACAGGCCACCGTATTACGGCTATGTCAAACGAGCTACTAACAGGGAACCCTCTGCTCATGACTATTGGTGG
GCTGAGCACCAGAAAACCTGTGGAGGCACTTACATAAAAATCAAGGAACCAGAGAATTACTCAAAAAAAGGCAAA
GGAAAGGCAAAACTAGGAAAGGAACCAGTATTGGCCGCAGAGAATAAAGGTACCTTCGTGTATATTCTTCTGATT
TTTATG**TGA**CCATAGCTATGATGTAAAGACAATACTGTCCTTCAGAGAACTGGTATTAAGATAAACTTAAGGATC
GTTTCTGGTGTAGAAGTCTTCAAGTGTAGACTTAAGGAAAAAATCCCACTGTCCATGAAATGATGGTAGGAAAAC
AGACTTTGCTCTGTACAGAAGTAAGTAAAAGTAGGAATAGTTTCCATGGATATTTTTATTTTTATTAACTTTTTT
CAGTTTCTTTTTATTCAAAGAAACAAAATTCAATCTCTGATAATATTTGAGGTAAAGTTCCTTTCCCTATCTTGA
CTCACTGAGTTATTAGGAAACAGAAGGCAAAAAGATTGTCAAAATAAAAACAATAATTCAAGTAACAATGCCCGG
AATATACGTCCTAACTACACCCCTTCCTATCAGCTGGATTCTATCCAAGTGACTCTATTGATGTATGTATGTTCA
TTCAAAGAATGGGAAAAGGATATGACATATATTTGCCAGTACTTCATCTTCAAGATTTACCCTTTTCCTGTGAAG
TTCAGAGTTACTGAAGATGCTTCTTCCCTTGGGAAGTTGTTGACCCAAGAACATAGGTTATATTTCCCAAATCTT
TAATTATTGAGTGAAAGAGCTATAGATGAATTGATATGGAAAGACCGTATCTTCATTTTCGTGAGTAGAAGGAAA
GATAAGAATGAGGCAGCAGATTTTCCCTCCTGGAATTACACATAAAGGACACTAAGCAATTTTCAAGGTAAATGT
TGCCTTGTTGTTGGTCTTTGGCATGATAAGATTCTTTATTTAAATATGAGAGAATTTTTTTTATCCTTTATATT
CTCTCAATATCAGAACTCCTGAATTCTGAAGATTGCCCTCCTCCCATTAATAGGATTGTATGGATGTAAGATGGA
ATAAAATACTAGTTCTTCATTTTGAGAAAACTGTACATTAGTTTAATGTTTGTTACTGTATTTCTTTTGAGTTGA
GGCACTTACATAACAATCTTCTTTGCTTTTTTGGCAGATAAACCCAACAGAGGTGAGGCCCAGCTAGTAATCCCT
TTTAGTGGGAAAGGATATGTTCTAGGAGAAACAAGCAATTTACCTTCACCTGGGAAACTGATCACTTCACATGCC
ATTAATAAAACCCAAGATCTTTTAAATCAAACCATTCAGCAAATGCTGTAAGACCTAATTCTAAAATCAAGGTG
AAATTTGAACAGAATGGTTCAAGTAAAAATTCTCATCTGGTCTCCCCTGCTGTTAGTAACAGTCACCAAAATGTT
CTAAGCAACTACTTTCCTAGAGTATCATTTGCCAACCAAAAGGCTTTCAGAGGTGTGAATGGATCTCCAAGGATA
AGTGTAACAGTTGGCAACATCCCTAAAAACTCAGTCTCTTCTAGTTCTCAGAGAAGGGTTTCATCTTCTAAGATA
TCCCTAAGAAATTCTTCAAAAGTAACGGAATCAGCATCTGTGATGCCATCCCAGGATGTGAGTGGGTCTGAAGAT
ACATTCCCAAATAAACGACCTAGGCTAGAAGATAAAAAAAAA
```

FIGURE 78

MDDDLMLALRLQEEWNLQEAERDHAQESLSLVDASWELVDPTPDLQALFVQFNDQFFWGQLEA
VEVKWSVRMTLCAGICSYEGKGGMCSIRLSEPLLKLRPRKDLVETLLHEMIHAYLFVTNNDKD
REGHGPEFCKHMHRINSLTGANITVYHTFHDEVDEYRRHWWRCNGPCQHRPPYYGYVKRATNR
EPSAHDYWWAEHQKTCGGTYIKIKEPENYSKKGKGKAKLGKEPVLAAENKGTFVYILLIFM

Important features of the protein:

Signal peptide:

amino acids 1-41

N-glycosylation sites.

amino acids 148-151, 217-220

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 184-187

Casein kinase II phosphorylation sites.

amino acids 30-33, 121-124, 154-157, 187-190, 192-195

Tyrosine kinase phosphorylation site.

amino acids 211-218

N-myristoylation sites.

amino acids 59-64, 85-90, 146-151

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 108-117

FIGURE 79

CGGACGCGTGGGTGGCAACCAGGAGAAGCCAAACTTGGTCCCCCGGCTCGCGGAGTGCCTGCG
AGCGGTGCTC**ATG**GCGCTCTATGAGGTCTTCTCTCACCCGGTCGAGCGCAGTTACCGCGCGGG
GCTCTGCTCCAAAGCCGCGCTGTTCCTGCTGCTGGCCGCTGCGCTCACGTACATCCCGCCGCT
GCTGGTGGCCTTCCGGAGCCACGGGTTTTGGCTGAAGCGGAGCAGCTACGAGGAGCAGCCGAC
CGTGCGCTTCCAACACCAGGTGCTGCTCGTGGCCCTGCTCGGACCCGAAAGCGACGGGTTCCT
CGCCTGGAGCACGTTCCCCGCCTTCAACCGGCTGCAAGGGGATCGCCTGCGCGTCCCGCTCGT
TTCGACTAGAGAAGAAGACAGGAACCAGGATGGGAAGACGGACATGTTACATTTTAAGCTGGA
GCTTCCCCTGCAGTCCACGGAGCACGTTCTCGGTGTGCAGCTCATCCTGACTTTCTCCTATCG
ATTACACAGGATGGCGACCCTCGTGATGCAGAGCATGGCGTTTCTCCAGTCCTCCTTTCCTGT
CCCGGGATCCCAGTTATACGTGAACGGAGACCTGAGGCTGCAGCAGAAGCAGCCGCTGAGCTG
TGGTGGCCTAGATGCCCGATACAACATATCCGTGATCAACGGGACCAGCCCCTTTGCCTATGA
CTACGACCTCACCCATATTGTTGCTGCCTACCAGGAGAGGAACGTTACCACCGTCCTGAATGA
TCCCAACCCCATCTGGCTGGTGGGCAGGGCCGCAGATGCTCCATTTGTGATTAATGCTATCAT
CCGATACCCTGTGGAAGTCATTTCTTATCAGCCAGGATTCTGGGAGATGGTAAAGTTCGCCTG
GGTACAGTATGTCAGCATCCTGCTTATCTTCCTCTGGGTGTTTGAAAGAATCAAGATCTTCGT
GTTTCAGAATCAGGTGGTGACCACCATTCCTGTGACAGTGACGCCCCGGGGAGACTTGTGTAA
GGAGCACTTATCC**TAG**AAAGGCCATTTCTGAAGACTCAGCAGGACCGTGGCTGCCTCATTGTC
ATCTTCTGGGAACATCTTAGGACCTTTTGAAAGAGCCCAGCGGACACCTGCGGGCTTGTGTGC
TTTTCCCTCAGAGACAACGGTTCTTTCCGGTTTTGCTCTACACAGTTCCGTATCTTCAGAGCT
CCTGCAGAATTGTCAGGGACTAGTTTGTGGAAAGGTCTGAGAGTTCCTGGAGGCTATAATTAG
CTTTTTGGGTTTTCCTTCTTTGCCTTAGCGTTGAATTTCAGGAGAAAATTGCAGTCAGTTCAG
ACATCTTGGAAAGAGTCCCATCTCTGGTCAAGCAGAGACTTTTCCTCTGTTGAACTGAGGAAC
ACACTGTGCATTTCTTCCTTCTGTTGTGAGCCACTCTTACTCTTTTCAGGGCTCTCTTGTGAC
AAACATGCCAATCACTAGCACTTTGCACCCCTGGGCTTCTCCATTTCCCATTCACAGCTTTGA
TTTCCAGAGCTGAGGCCTTTAACTGGAGACCTGGAGGGGCAGGGCCCAAGGGCAAGGGCCGCA
TTAGCACAGGCAATCAGGGAGGGCCGCTGAAGGACACTTGGACCGTCCACCTGCCCCAGCCCA
ACAGTCAGTCATCTGTCATCAGCTCAGCTGAGCAGCCCTGGATCTTTGCCGTACTGTGACTGG
GCTCTTTGCCCTATTTTTCCCTCTGTCTGTGCCCCTGGATGGCAGGCTGAAGTCAGAGGGGCT
GTTTCATTCTCAGCCCCCTCAGCAGCACTGGGGGAAGAAAGCATTGTCACAACAGGTTCTTTC
TGGCCCTCACCCAACAGCCTGGGCACTTGGCCCTCCTCCTCCTTGACAGCCCTCCCCCTTCCT
GCAAAGGACAGGGGCGACAGGGGTTGGTGTTGGGATTGGCTCCCGCTGCCTGACAACCACAAG
TTTATTTGGAAGGCTAGCGGGAAGCCCAGCGGCTGGCGTTTCCCTTGACTAAGGAACAGGGTG
CCCATCAGAGTGGGGCGGGCAGCTTTGGGAAGGACACAAGAAGCAGTAAGAGTGTAAAGAGGA
TGCTGGCCTGGGCAGGCCAGTCCAGCCTGGCCACTAGCAGAATACCAAGCAGTCCAGTGGATT
ACCCTCGTGGCTAAGCAAGTGTCTGCAGGAGCAGAGATGGCTGGAAGGGGCCTCTGCACACGG
AAGATGGCTTGTTCAGCCCATTCACCTCCTGAGGATGTGGGCAGTCTCCTCCAAGAACACATG
GAGCTGCTTCCTGATCCCAAGCAGGTCATTGCCACTGGAAGGACATGGCCCCGGTGATCCATG
CTTCATGCCCACCCAGAAACACACCCCTCAGTGTGTGCCTCAGTTTACTTTGGAGATCAGTTG
TCGTTTTTAGTGCTCCTTTAGGCTTACTAAAACAGTTTTGGAAACAAAGCTATTTTGAAGTAT
TCAAGCAGAGGAATTCCCTAACACTGACCCCCTTGTCTTTTTTTAATATTCAGGCTGTTTTAT
ATGCCTAAATTTTTTTCTTAAGATCTAAACGAAAAATAGTTTCTTGTTTAAATTCACATAAGG
CAATGAGATATGGAAAGATGACAAGATACGTATAAACATTGGTTTGCATCTTATTAAATTATT
CTAATGCAAATCTTGTATAAAGAACCCATGATGTTTTGTAACTTTCTAATTAAAATGTTCAAA
ATGAG

FIGURE 80

MALYEVFSHPVERSYRAGLCSKAALFLLLAAALTYIPPLLVAFRSHGFWLKRSSYEEQPTVRF
QHQVLLVALLGPESDGFLAWSTFPAFNRLQGDRLRVPLVSTREEDRNQDGKTDMLHFKLELPL
QSTEHVLGVQLILTFSYRLHRMATLVMQSMAFLQSSFPVPGSQLYVNGDLRLQQKQPLSCGGL
DARYNISVINGTSPFAYDYDLTHIVAAYQERNVTTVLNDPNPIWLVGRAADAPFVINAIIRYP
VEVISYQPGFWEMVKFAWVQYVSILLIFLWVFERIKIFVFQNQVVTTIPVTVTPRGDLCKEHLS

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 268-284

N-glycosylation sites.

amino acids 194-198, 199-203, 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 51-55

Tyrosine kinase phosphorylation site.

amino acids 250-259

N-myristoylation site.

amino acids 187-193

Cell attachment sequence.

amino acids 307-310

FIGURE 81

```
GCCGGGAGCTTCCCTGATGGTGCCGCCGCCTCCGAGCCGGGGAGGAGCTGCCAGGGGCCAGCTGGGCAGGAGCCT
GGGTCCGCTGCTGCTGCTCCTGGCGTTGGGACACACGTGGACCTACAGAGAGGAGCCGGAGGACGGCGACAGAGA
AATCTGCTCAGAGAGCAAAATCGCGACGACTAAATACCCGTGTCTGAAGTCTTCAGGCGAGCTCACCACATGCTA
CAGGAAAAAGTGCTGCAAAGGATATAAATTTGTTCTTGGACAATGCATCCCAGAAGATTACGACGTTTGTGCCGA
GGCTCCCTGTGAACAGCAGTGCACGGACAACTTTGGCCGAGTGCTGTGTACTTGTTATCCGGGATACCGATATGA
CCGGGAGAGACACCGGAAGCGGGAGAAGCCATACTGTCTGGATATTGATGAGTGTGCCAGCAGCAATGGGACGCT
GTGTGCCCACATCTGCATCAATACCTTGGGCAGCTACCGCTGCGAGTGCCGGGAAGGCTACATCCGGGAAGATGA
TGGGAAGACATGTACCAGGGGAGACAAATATCCCAATGACACTGGCCATGAGAAGTCTGAGAACATGGTGAAAGC
CGGAACTTGCTGTGCCACATGCAAGGAGTTCTACCAGATGAAGCAGACCGTGCTGCAGCTGAAGCAAAAGATTGC
TCTGCTCCCCAACAATGCAGCTGACCTGGGCAAGTATATCACTGGTGACAAGGTGCTGGCCTCAAACACCTACCT
TCCAGGACCTCCTGGCCTGCCTGGGGGCCAGGGCCCTCCCGGCTCACCAGGACCAAAGGGAAGCCCAGGCTTCCC
CGGTATGCCAGGCCCTCCTGGGCAGCCCGGCCCACGGGGCTCAATGGGACCCATGGGACCATCTCCTGATCTGTC
CCACATTAAGCAAGGCCGGAGGGGCCCTGTGGGTCCACCAGGGGCACCAGGAAGAGATGGTTCTAAGGGGGAGAG
AGGAGCGCCTGGGCCCAGAGGGTCTCCAGGACCCCCTGGTTCTTTCGACTTCCTGCTACTTATGCTGGCTGACAT
CCGCAATGACATCACTGAGCTGCAGGAAAAGGTGTTCGGGCACCGGACTCACTCTTCAGCAGAGGAGTTCCCTTT
ACCTCAGGAATTTCCCAGCTACCCAGAAGCCATGGACCTGGGCTCTGGAGATGACCATCCAAGAAGAACTGAGAC
AAGAGACTTGAGAGCCCCCAGAGACTTCTACCCATAGCACATCCCAACACCGTCACGCCAAAGGAAGAGAAAGAT
CAACTCACCTGCAGTTAAACCATCTAAAGAGAAGAAAGACCACTGGAGACCTAGAAAACATACATTTTTCTCTTC
TCTTCTCCTGACGTCTCTCCACTCCTCTTCTTCCAAATACGATGCTATTTTCAGAGTCCCCTCCTAGGCCTGCAG
ACATGAGGGAGTGAATGATTGATTTACCTGCTTCTCACTAAGAGTCCATTGGGGTGGTTTGCATTGTAACTTTTC
TTTTACATCCTATTTTTCCAGGAACTTTGGATTTAAGTACTCTCACAGTGTCTTAAATCATAAATTCTTGAAGTT
AAATTTGGCAGAGTATCAAAAGGGGGAAAATGACAAAGTGAGCTCTAAGAAAATGTGAGGCTACTTCTAAGATGT
GTGTTCACAATAGACCATAACTCCTCTAGTATCAAAATTGGGGCTCTTCAGTTAAAAAGGGGTGGGGAGGACAAA
CGTGTCGATGTGCTTTGGTGGAGAATTTTTTCCTTGTGCTTCTAGTAGACTTTAAATATTGTATCCCTTTGTCAA
ACCTTGTTTCCCAAATTCAATTAAAGAGAGGAGAGAATTGAATGGCGTTTAGAGAAGATAGAAAAGAATCACAGT
CATATATTTACTGTTATATAGATTGCCACATTCTAAAATTCAAATACGGTGCTTAAGGTTTCATGCCATGCTTAT
CTGTAAGTATCCTATTTAGGGAAGAAGATTAAACTCTCTTTTCAAAAAAACAAAGTGAAATGCCTGGATTCACAT
TAAAACAATGGGCTCTCGTTTGCTATAATATTTTAAAGCTGTTTAATCAACAGTGGAGTCTGCTCTATAAATATA
GATTATTTGTTCAATAAACTGGCTGAGCTTAGAGAGAGGTGCAGAATTCCTGGTTCTGAGCAGGTGCCCAGAAGG
TACCATTAGGTGCCATGATCCAGGCTGAACCAATATACAGTGGGGCTGAAGTCTGCAAGGAGGTTGCTGGCTTGG
GCTGACCTCACTAATGCCATCAGCAGCGGTAGGTAAATTTTTTCTCCTTGGGTATTACAAGTTTTTGTCTGGAGC
CAACCAAGCTTGCCACCAACATATTGAGAGTAATACACTATTGAAAGTTATCTTGGATGGGGAGAAAAAAAAATA
GTGGTTTTCCTTGTTTGCAAAAACTTCCTTCCTATTCTCATTTTTTCTTAATTTTCTTTAATTTAGTCCAAGTTC
CAGTTCTTTTAGGCCTTCTCTTTGATTTATTTTCCCCTGCATGTGAGAAGCAGTTCAGAAAAAGGTCTATATCTC
CACCTCCTAGTGAGTTAGAGTGTTTTCTCAGAGCACCTCTGGGTGGCAAAGGGAAGCATGTTCCTGCCAAGGTTT
GCTGTGGATTCAGAAGCACCAGGACAAGAGACCAGAAGGATGATCTGCTCCTTTGTAACGTTGTTGAGGGCCCT
CTTGTTTCCAATGAGCAGCTTATAGGTTACTCACAGTCCACTTTCTCACTGGACACACAAAGTGGCTCTTTATCT
ACCTTTGCGGGAGATTTTCACTCTCCTGCAAATGATCGTTCTCACACTCATATTAGCTCATGTTGGAATTTCCCA
TCCTGCCATGTCCTTTCCCATTTCTTTTTGGCTTTTTTGCCTCCACCTTTTAGCCCACATCATTTAACTCCACTA
CTGTGAAAGCTTGCTTAAAGAAAATCCCTCTTGGCCGGGTGTGGTAGCCCACGCCTCTAATCCCAGCACTTTGGG
AGGCTGAGGCGGGGAGATCACAAGGTCAGGAGATCGAGACCAGCCTGACCAACATGGTGAAACCCTGTCTCTACT
AAAAATACAAAAATTAGCTGGGCGTGTTGGCACACACCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAA
TTACTTTAACCTGCGGGGGGAGCCTAGATTGCGCTACTGCACTCCAGCCTAGGCAACAGAGGGAGACTCTGTCTC
ATTAAAAA
```

FIGURE 82

MVPPPPSRGGAARGQLGRSLGPLLLLLALGHTWTYREEPEDGDREICSESKIATTKYPCLKSS
GELTTCYRKKCCKGYKFVLGQCIPEDYDVCAEAPCEQQCTDNFGRVLCTCYPGYRYDRERHRK
REKPYCLDIDECASSNGTLCAHICINTLGSYRCECREGYIREDDGKTCTRGDKYPNDTGHEKS
ENMVKAGTCCATCKEFYQMKQTVLQLKQKIALLPNNAADLGKYITGDKVLASNTYLPGPPGLP
GGQGPPGSPGPKGSPGFPGMPGPPGQPGPRGSMGPMGPSPDLSHIKQGRRGPVGPPGAPGRDG
SKGERGAPGPRGSPGPPGSFDFLLLMLADIRNDITELQEKVFGHRTHSSAEEFPLPQEFPSYP
EAMDLGSGDDHPRRTETRDLRAPRDFYP

Important features of the protein:

Signal peptide:

amino acids 1-34

N-glycosylation sites.

amino acids 142-148, 182-188

Tyrosine kinase phosphorylation site.

amino acids 125-132

N-myristoylation sites.

amino acids 10-16, 143-149, 155-161, 196-202, 250-256

Amidation site.

amino acids 299-303

Aspartic acid and asparagine hydroxylation site.

amino acids 150-162

Cell attachment sequence.

amino acids 176-179

C1q domain proteins.

amino acids 247-280

Calcium-binding EGF-like domain proteins pattern proteins.

amino acids 144-165

FIGURE 83

ATCTGAGTGAGCTAACTGACACA**ATG**AAACTGTCAGGCATGTTTCTGCTCCTCTCTCTGGCTC
TTTTCTGCTTTTTAACAGGTGTCTTCAGTCAGGGAGGACAGGTTGACTGTGGTGAGTTCCAGG
ACCCCAAGGTCTACTGCACTCGGGAATCTAACCCACACTGTGGCTCTGATGGCCAGACATATG
GCAATAAATGTGCCTTCTGTAAGGCCATAGTGAAAAGTGGTGGAAAGATTAGCCTAAAGCATC
CTGGAAAATGC**TGA**GTTAAAGCCAATGTTTCTTGGTGACTTGCCAGCTTTTGCAGCCTTCTTT
TCTCACTTCTGCTTATACTTTTGCTGGTGGATTCCTTTAATTCATAAAGACATACCTACTCTG
CCTGGGTCTTGAGGAGTTCAATGTATGTCTATTTCTCTTGATTCACTTGTCAATAAAGTACATTC
TGCAAAAGCAAAAA

FIGURE 84

MKLSGMFLLLSLALFCFLTGVFSQGGQVDCGEFQDPKVYCTRESNPHCGSDGQTYGNKCAFCK
AIVKSGGKISLKHPGKC

Important features of the protein:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 26-32, 52-58, 56-62, 69-75

Kazal serine protease inhibitors family signature.

amino acids 40-63

FIGURE 85

```
GGAGCAGACACACAGACCCGGGCCGGAGGCCCCTCTTCTAGCCCTGCGGGAACCGGACAGTTC
CCCAACTGGGGACTCTGGAACCACAGCTCCTAAATCATCAAATTCTCAAGCTTTTTTTTTCCC
TCTCTTCGTCCCAGCCATCCCAGTCTTCTTCTTCTTTTTTTTTTTTTTAACTTATTGTTTTTT
TCGCTCCTGTCATTATGAAAGTGGTCACGCCATTCAATATTAAGACTTGGAGGGAATTGGGGA
AAGAAAAGAAAGAATCTAAAAGAAGAGAAGCGACCGGTGCTTTTAAGGGTGTCTAATTTTCAA
AAGAGACGTCTGGGAGTATTTTGCTCTGGGCGTTTGGAGCAACTTCGCGGACAGCGGAGCTCG
CCCAGCATGGATGTTCCAGGTTCACAGGCGCCTTTCTTCTGAGAACGACCCTGGCCTTGAACG
TCAGAGCCGGGGACGAAGGCCCCCGGAGGCTGCTGCGAGCTCCGCGCGTTCCTTCGCGCCCTT
CCGCGCCGCTCGCGCCGGCGCCGGCCTCCACCCCCGCGCGCCGCCTCCCACCAGTCCCG**ATG**C
AGGCGCCCGGCCGGGGGCCACTCGGGCTGCGGCTGATGATGCCCGGGCGCCGGGGGGCGCTGC
GCGAGCCTGGCGGCTGCGGATCCTGCCTGGGGGTGGCGCTGGCCCTGCTGTTGCTGCTACTGC
CCGCCTGCTGCCCCGTGCGGGCGCAGAACGACACGGAGCCCATCGTGCTGGAGGGCAAGTGCC
TGGTGGTGTGCGACTCCAGCCCGTCGGCGGACGGCGCCGTCACCTCCTCCCTAGGCATCTCCG
TGCGCTCCGGCAGCGCCAAGGTGGCCTTCTCCGCCACGCGGAGCACCAACCACGAGCCGTCCG
AGATGAGCAACCGCACCATGACCATCTATTTCGACCAGGTATTAGTAAATATTGGCAACCACT
TTGATCTTGCTTCCAGTATATTTGTAGCACCGAGAAAAGGGATTTATAGCTTCAGCTTCCACG
TGGTCAAAGTGTATAACAGACAAACCATCCAGGTCAGTTTAATGCAGAATGGCTACCCAGTGA
TCTCGGCCTTTGCAGGAGACCAGGATGTCACCAGAGAAGCTGCTAGCAATGGCGTGCTGCTGC
TCATGGAAAGGGAAGACAAAGTGCATCTCAAACTTGAGAGAGGCAACCTCATGGGGGGCTGGA
AATACTCCACATTCTCGGGCTTCTTGGTGTTTCCTCTA**TAA**ACACAGAGCCCCCTAGATGGTG
GGGGAATGGCAAACTGGACCCAGGACTCCGCCCTTTAAAACACCCTGAACTTACTGGAATTGG
ACACCTTGTTTCCAACCTCCGTCAGACTGTTGCAGTAGAAGAATGATTTCCTTTGAAACCTCC
AGTACTTTTGTTTTTGTTTTTTGGAATACTGACAATTCCTCGGGAACCTGGCCTCTAATTAGT
TTTAGATGACAAGGTCTTAAGGAGAAATGAAATTATCGATTTGAGCAATTTGTACCTGTGATT
GTAAAGTCAATATCGGATTTTATTGTTGGGACCATGGACCTCTTTTGTTTGTATGTTGTATTG
TCGTCCCAACGGAAGGAGAGCTCCTGACTCCAGGATGGGCTGCAGGTTGCAGTCAGGGCTTGA
AGTAGGAGCCCAGCAAAGAACCACCTGCTGGACAGTCCTTGACATGTGTTCTGTGTGTGTCTG
TATAGCCTTAAGAAAAGAATGGCTTCACTTTCATTCTGTATTCTTCCCCCCACCATGTGGCT
GGGAGGACTTGGGAGGGGGATGGGGACATTGGGAACCTGTCAAGAAGTGCTTTATCCAGAGAA
GCAAATTTTGCACGATTGGACTGCAATTTTTGTTTTGTATTGTTTGTGTTTTTTCTTGAAAAG
CTTTACTTTTCTTTCCACACTCAGCTCTCCCTCCTCAACCCCACTTTTATTTTTCTTGCTGGG
GTTGAGGAGAGAAAATATAGAATTCCTGGATAAGACCAAACAAAACAAAACATTAAAATACCT
GTATGTTTTGTTTTAGACGAGACCAAACTAAACAAAAAGTATCTGTTTATCAAAGTAAAAGTA
ACACAATGGACAATTCTGCTTATTCTCTCAAAGAGATTCTAAGATGCACCTTTAGAACTATTA
ATAGCAACCTGCATTTTTTTTTAATTTATACTTCAGAATCCTTTAAGAACCTGGTGTTCCTGA
GTGGTCCTGAATCATATAAGTTGGTAATGGAAGCTGTAATGACCAAGTCCCCTAAACATACTA
TGTCTTTGCCACGTGTGCTGTGACTTCTCTGTGGGTGATTTAATTTATTTGGATCCACCTCTG
AGTGAGCGCACAGTGATCAGGTGCTTCAAAGCCAACAGACCAGCTCCTCTTCCTCCGGATCCT
CTTTTGATCTGCCCAGGAAAGGGATGCATTGACACTCTCCTGCATGCACCTGGCGAGAAGCCA
CCTGAAAGTCACTGTGGTTAAAGATATTGGTGGAGGTACCCCAGGAGCACTGTTACAAATCCT
TCTTGTTTTGGCATCTCGTACAACATTATTAAGACACAGCTGAGAGTTGATGGGTGTGTAATG
CATATGCCAAGGAAATGTCACTAATCCCAAAGCAATCAAAAAGGAGACCTCAAACCAGATGTT
AATTTGTTCTTTGTGTAACAATGTAACCAAAATATTGATGATAAAAGTCATAATTTAAGATTC
AGAATAAATGGGTTTGATGTCTGGCAAAAAAAAAAAAAAAA
```

FIGURE 86

MQAPGRGPLGLRLMMPGRRGALREPGGCGSCLGVALALLLLLLPACCPVRAQNDTEPIVLEGK
CLVVCDSSPSADGAVTSSLGISVRSGSAKVAFSATRSTNHEPSEMSNRTMTIYFDQVLVNIGN
HFDLASSIFVAPRKGIYSFSFHVVKVYNRQTIQVSLMQNGYPVISAFAGDQDVTREAASNGVL
LLMEREDKVHLKLERGNLMGGWKYSTFSGFLVFPL

Important features of the protein:

Signal peptide:

amino acids 1-48

N-glycosylation sites.

amino acids 53-57, 110-114

N-myristoylation sites.

amino acids 26-32, 27-33, 29-35, 33-39, 76-82, 205-211

Amidation site.

amino acids 16-20

C1q domain signature.

amino acids 117-148

C1q domain proteins.

amino acids 115-149

FIGURE 87

AGGGCCCGCGGGTGGAGAGAGCGACGCCCGAGGGG**ATG**GCGGCAGCGTCCCGGAGCGCCTCTG
GCTGGGCGCTACTGCTGCTGGTGGCACTTTGGCAGCAGCGCGCGGCCGGCTCCGGCGTCTTCC
AGCTGCAGCTGCAGGAGTTCATCAACGAGCGCGGCGTACTGGCCAGTGGGCGGCCTTGCGAGC
CCGGCTGCCGGACTTTCTTCCGCGTCTGCCTTAAGCACTTCCAGGCGGTCGTCTCGCCCGGAC
CCTGCACCTTCGGGACCGTCTCCACGCCGGTATTGGGCACCAACTCCTTCGCTGTCCGGGACG
ACAGTAGCGGCGGGGGGCGCAACCCTCTCCAACTGCCCTTCAATTTCACCTGGCCGGGTACCT
TCTCGCTCATCATCGAAGCTTGGCACGCGCCAGGAGACGACCTGCGGCCAGAGGCCTTGCCAC
CAGATGCACTCATCAGCAAGATCGCCATCCAGGGCTCCCTAGCTGTGGGTCAGAACTGGTTAT
TGGATGAGCAAACCAGCACCCTCACAAGGCTGCGCTACTCTTACCGGGTCATCTGCAGTGACA
ACTACTATGGAGACAACTGCTCCCGCCTGTGCAAGAAGCGCAATGACCACTTCGGCCACTATG
TGTGCCAGCCAGATGGCAACTTGTCCTGCCTGCCCGGTTGGACTGGGGAATATTGCCAACAGC
CTATCTGTCTTTCGGGCTGTCATGAACAGAATGGCTACTGCAGCAAGCCAGCAGAGTGCCTCT
GCCGCCCAGGCTGGCAGGGCCGGCTGTGTAACGAATGCATCCCCCACAATGGCTGTCGCCACG
GCACCTGCAGCACTCCCTGGCAATGTACTTGTGATGAGGGCTGGGGAGGCCTGTTTTGTGACC
AAGATCTCAACTACTGCACCCACCACTCCCCATGCAAGAATGGGGCAACGTGCTCCAACAGTG
GGCAGCGAAGCTACACCTGCACCTGTCGCCCAGGCTACACTGGTGTGGACTGTGAGCTGGAGC
TCAGCGAGTGTGACAGCAACCCCTGTCGCAATGGAGGCAGCTGTAAGGACCAGGAGGATGGCT
ACCACTGCCTGTGTCCTCCGGGCTACTATGGCCTGCACTGTGAACACAGCACCTTGAGCTGCG
CCGACTCCCCCTGCTTCAATGGGGGCTCCTGCCGGGAGCGCAACCAGGGGGCCAACTATGCTT
GTGAATGTCCCCCCAACTTCACCGGCTCCAACTGCGAGAAGAAAGTGGACAGGTGCACCAGCA
ACCCCTGTGCCAACGGGGGACAGTGCCTGAACCGAGGTCCAAGCCGCATGTGCCGCTGCCGTC
CTGGATTCACGGGCACCTACTGTGAACTCCACGTCAGCGACTGTGCCCGTAACCCTTGCGCCC
ACGGTGGCACTTGCCATGACCTGGAGAATGGGCTCATGTGCACCTGCCCTGCCGGCTTCTCTG
GCCGACGCTGTGAGGTGCGGACATCCATCGATGCCTGTGCCTCGAGTCCCTGCTTCAACAGGG
CCACCTGCTACACCGACCTCTCCACAGACACCTTTGTGTGCAACTGCCCTTATGGCTTTGTGG
GCAGCCGCTGCGAGTTCCCCGTGGGCTTGCCGCCCAGCTTCCCCTGGGTGGCCGTCTCGCTGG
GTGTGGGGCTGGCAGTGCTGCTGGTACTGCTGGGCATGGTGGCAGTGGCTGTGCGGCAGCTGC
GGCTTCGACGGCCGGACGACGGCAGCAGGGAAGCCATGAACAACTTGTCGGACTTCCAGAAGG
ACAACCTGATTCCTGCCGCCCAGCTTAAAAACACAAACCAGAAGAAGGAGCTGGAAGTGGACT
GTGGCCTGGACAAGTCCAACTGTGGCAAACAGCAAAACCACACATTGGACTATAATCTGGCCC
CAGGGCCCCTGGGGCGGGGGACCATGCCAGGAAAGTTTCCCCACAGTGACAAGAGCTTAGGAG
AGAAGGCGCCACTGCGGTTACACAGTGAAAAGCCAGAGTGTCGGATATCAGCGATATGCTCCC
CCAGGGACTCCATGTACCAGTCTGTGTGTTTGATATCAGAGGAGAGGAATGAATGTGTCATTG
CCACGGAGGTA**TAA**GGCAGGAGCCTACCTGGACATCCCTGCTCAGCCCCGCGGCTGGACCTTC
CTTCTGCATTGTTTACA

FIGURE 88

MAAASRSASGWALLLLVALWQQRAAGSGVFQLQLQEFINERGVLASGRPCEPGCRTFFRVCLK
HFQAVVSPGPCTFGTVSTPVLGTNSFAVRDDSSGGGRNPLQLPFNFTWPGTFSLIIEAWHAPG
DDLRPEALPPDALISKIAIQGSLAVGQNWLLDEQTSTLTRLRYSYRVICSDNYYGDNCSRLCK
KRNDHFGHYVCQPDGNLSCLPGWTGEYCQQPICLSGCHEQNGYCSKPAECLCRPGWQGRLCNE
CIPHNGCRHGTCSTPWQCTCDEGWGGLFCDQDLNYCTHHSPCKNGATCSNSGQRSYTCTCRPG
YTGVDCELELSECDSNPCRNGGSCKDQEDGYHCLCPPGYYGLHCEHSTLSCADSPCFNGGSCR
ERNQGANYACECPPNFTGSNCEKKVDRCTSNPCANGGQCLNRGPSRMCRCRPGFTGTYCELHV
SDCARNPCAHGGTCHDLENGLMCTCPAGFSGRRCEVRTSIDACASSPCFNRATCYTDLSTDTF
VCNCPYGFVGSRCEFPVGLPPSFPWVAVSLGVGLAVLLVLLGMVAVAVRQLRLRRPDDGSREA
MNNLSDFQKDNLIPAAQLKNTNQKKELEVDCGLDKSNCGKQQNHTLDYNLAPGPLGRGTMPGK
FPHSDKSLGEKAPLRLHSEKPECRISAICSPRDSMYQSVCLISEERNECVIATEV

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 530-552

N-glycosylation sites.

amino acids 108-112, 183-187, 205-209, 393-397, 570-574, 610-614

Glycosaminoglycan attachment site.

amino acids 96-100

Tyrosine kinase phosphorylation site.

amino acids 340-347

N-myristoylation sites.

amino acids 42-48, 204-210, 258-264, 277-283, 297-303, 383-389,
415-421, 461-467, 522-528, 535-541, 563-569, 599-605, 625-631

Amidation site.

amino acids 471-475

Aspartic acid and asparagine hydroxylation site.

amino acids 339-351

EGF-like domain cysteine pattern signature.

amino acids 173-185, 206-218, 239-251, 270-282, 310-322, 348-360,
388-400, 426-438, 464-476, 506-518

Calcium-binding EGF-like:

amino acids 224-245, 255-276, 295-316, 333-354, 373-394, 411-432,
449-470

FIGURE 89

GTCTCCGCGTCACAGGAACTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGAC

TTGACTCCCGCGCGCCCCAACCCTGCTTATCCCTTGACCGTCGAGTGTCAGAGATCCTGCAGC

CGCCCAGTCCCGGCCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCCTGTTTTTACTCC

TCCTTTTCATTCATAACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCC

GAGCGTGGAAGA**ATG**GGGTTCCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCCCGATT

CAAGCTTTCCCCAAACCTGGAGGAAGCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCA

GAAAGACCTTTGAATGAACAGATTGCTGAAGCAGAAGAAGACAAGATTAAAAAAACATATCCT

CCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTTTGTTGATAACTTGAACCTGCTAAAGGCA

ATAACAGAAAAGGAAAAAATTGAGAAAGAAGACAATCTATAAGAAGCTCCCCACTTGATAAT

AAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACTGATCGATGATTATGACTCT

ACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATCAACTAGACGGG

ACTCCTTTAACCGCTGAAGACATTGTCCATAAATCGCTGCCAGGATTTATGAAGAAAATGAC

AGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGCCAA

GCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAAC

AATTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCA

GAGAAAGTGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACA

GTATCTAACACATTAACCTTGACAAATGGCTTGGAAAGGAGAACTAAAACCTACAGTGAAGAC

AACTTTGAGGAACTCCAATATTTCCCAAATTTCTATGCGCTACTGAAAGTATTGATTCAGAA

AAAGAAGCAAAAGAGAAAGAAACACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAG

ATGATGGTGAAATATGGAACAATATCTCCAGAAGAAGGTGTTTCCTACCTTGAAAACTTGGAT

GAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAAAAATGCTACTGACAATATAAGCAAG

CTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAACAGACAGTACCAAGGAAGAAGCAGCT

AAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAAGATGATAACTCCAACCCAGGA

GGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGCCATCAGAAAAAATATT

GAATGGTTGAAGAAACATGACAAAAAGGGAAATAAAGAAGATTATGACCTTTCAAAGATGAGA

GACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAAGAAGCC

GAGGCCATCAAGCGCATTTATAGCAGCCTG**TAA**AAATGGCAAAGATCCAGGAGTCTTTCAAC

TGTTTCAGAAAACATAATATAGCTTAAAACACTTCTAATTCTGTGATTAAAATTTTTTGACCC

AAGGGTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAAACATAGC

TTTCTTCCCGTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAA

AAA

FIGURE 90

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENK

PGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSG

LDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTL

EDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTPMAAIQDGLAKGENDETVSNT

LTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVK

YGTISPEEGVSYLENLDEMIALQTKNKLEKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEK

EYGSLKDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFIN

KQADAYVEKGILDKEEAEAIKRIYSSL

Important features:

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 91

TGCATCAGTGCCCAGGCAAGCCCAGGAGTTGACATTTCTCTGCCCAGCC**ATG**GGCCTCACCCT

GCTCTTGCTGCTGCTCCTGGGACTAGAAGGTCAGGGCATAGTTGGCAGCCTCCCTGAGGTGCT

GCAGGCACCCGTGGGAAGCTCCATTCTGGTGCAGTGCCACTACAGGCTCCAGGATGTCAAAGC

TCAGAAGGTGTGGTGCCGGTTCTTGCCGGAGGGGTGCCAGCCCCTGGTGTCCTCAGCTGTGGA

TCGCAGAGCTCCAGCGGGCAGGCGTACGTTTCTCACAGACCTGGGTGGGGGCCTGCTGCAGGT

GGAAATGGTTACCCTGCAGGAAGAGGATGCTGGCGAGTATGGCTGCATGGTGGATGGGGCCAG

GGGGCCCCAGATTTTGCACAGAGTCTCTCTGAACATACTGCCCCCAGAGGAAGAAGAAGAGAC

CCATAAGATTGGCAGTCTGGCTGAGAACGCATTCTCAGACCCTGCAGGCAGTGCCAACCCTTT

GGAACCCAGCCAGGATGAGAAGAGCATCCCCTTGATCTGGGGTGCTGTGCTCCTGGTAGGTCT

GCTGGTGGCAGCGGTGGTGCTGTTTGCTGTGATGGCCAAGAGGAAACAAGAATCCCTCCTCAG

TGGTCCACCACGTCAG**TGA**CTCTGGACCGGCTGCTGAATTGCCTTTGGATGTACCACACATTA

GGCTTGACTCACCACCTTCATTTGACAATACCACCTACACCAGCCTACCTCTTGATTCCCCAT

CAGGAAAACCTTCACTCCCAGCTCCATCCTCATTGCCCCCTCTACCTCCTAAGGTCCTGGTCT

GCTCCAAGCCTGTGACATATGCCACAGTAATCTTCCCGGGAGGGAACAAGGGTGGAGGGACCT

CGTGTGGGCCAGCCCAGAATCCACCTAACAATCAGACTCCATCCAGCTAAGCTGCTCATCACA

CTTTAAACTCATGAGGACCATCCCTAGGGGTTCTGTGCATCCATCCAGCCAGCTCATGCCCTA

GGATCCTTAGGATATCTGAGCAACCAGGGACTTTAAGATCTAATCCAATGTCCTAACTTTACT

AGGGAAAGTGACGCTCAGACATGACTGAGATGTCTTGGGGAAGACCTCCCTGCACCCAACTCC

CCCACTGGTTCTTCTACCATTACACACTGGGCTAAATAAACCCTAATAATGATGTGCAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 92

MGLTLLLLLLLGLEGQGIVGSLPEVLQAPVGSSILVQCHYRLQDVKAQKVWCRFLPEGCQPLV

SSAVDRRAPAGRRTFLTDLGGGLLQVEMVTLQEEDAGEYGCMVDGARGPQILHRVSLNILPPE

EEEETHKIGSLAENAFSDPAGSANPLEPSQDEKSIPLIWGAVLLVGLLVAAVVLFAVMAKRKQ

ESLLSGPPRQ

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 161-181

N-myristoylation sites.

amino acids 17-23, 172-178

Amidation site.

amino acids 73-79

FIGURE 93

```
GGCGGCGTTGCCGGGCTCTCCGGAAGGAGACGTGGCGGCGGTTGGGCCGGTGATACCCGGGCG
CTTTATAGTCCCGCCGCCTCCTCCTCCACCTCCTCCTCCTCCTCTCCTCCTGGGGCAGAG
GAGGTTGTGGCGGTGGCTGGAGAAAGCGGCGGCGGAGGATGGAGGAAGGAGGCGGCGGCGTAC
GGAGTCTGGTCCCGGGCGGGCCGGTGTTACTGGTCCTCTGCGGCCTCCTGGAGGCGTCCGGCG
GCGGCCGAGCCCTTCCTCAACTCAGCGATGACATCCCTTTCCGAGTCAACTGGCCCGGCACCG
AGTTCTCTCTGCCCACAACTGGAGTTTTATATAAAGAAGATAATTATGTCATCATGACAACTG
CACATAAAGAAAATATAAATGCATACTTCCCCTTGTGACAAGTGGGGATGAGGAAGAAGAAA
AGGATTATAAAGGCCCTAATCCAAGAGAGCTTTTGGAGCCACTATTTAAACAAAGCAGTTGTT
CCTACAGAATTGAGTCTTATTGGACTTACGAAGTATGTCATGGAAACACATTCGGCAGTACC
ATGAAGAGAAAGAAACTGGTCAGAAAATAAATATTCACGAGTACTACCTTGGGAATATGTTGG
CCAAGAACCTTCTATTTGAAAAAGAACGAGAAGCAGAAGAAAGGAAAAATCAAATGAGATTC
CCACTAAAAATATCGAAGGTCAGATGACACCATACTATCCTGTGGGAATGGGAAATGGTACAC
CTTGTAGTTTGAAACAGAACCGGCCCAGATCAAGTACTGTGATGTACATATGTCATCCTGAAT
CTAAGCATGAAATTCTTTCAGTAGCTGAAGTTACAACTTGTGAATATGAAGTTGTCATTTTGA
CACCACTCTTGTGCAGTCATCCTAAATATAGGTTCAGAGCATCTCCTGTGAATGACATATTTT
GTCAATCACTGCCAGGATCTCCATTTAAGCCCCTCACCCTGAGGCAGCTGGAGCAGCAGGAAG
AAATACTAAGGGTGCCTTTTAGGAGAAATAAAGAGGGTGTCGGTTGGTGGAAATATGAATTCT
GCTATGGCAAACATGTACATCAATACCATGAGGACAAGGATAGTGGGAAAACCTCTGTGGTTG
TCGGGACATGGAACCAAGAAGAGCATATTGAATGGGCTAAGAAGAATACTGCTAGAGCTTATC
ATCTTCAAGACGATGGTACCCAGACAGTCAGGATGGTGTCACATTTTTATGGAAATGGAGATA
TTTGTGATATAACTGACAAACCAAGACAGGTGACTGTAAAACTAAAGTGCAAAGAATCAGATT
CACCTCATGCTGTTACTGTATATATGCTAGAGCCTCACTCCTGTCAATATATTCTTGGGGTTG
AATCTCCAGTGATCTGTAAAATCTTAGATACAGCAGATGAAAATGGACTTCTTTCTCTCCCCA
ACTAAAGGATATTAAAGTTAGGGGAAAGAAAAGATCATTGAAAGTCATGATAATTTCTGTCCC
ACTGTGTCTCATTATAGAGTTCTCAGCCATTGGACCTCTTCTAAAGGATGGTATAAAATGACT
CTCAACCACTTTGTGAATACATATGTGTATATAAGAGGTTATTGATAAACTTCTGAGGCAGAC
ATTTGTCTCGCTTTTTTTCATTTTTGTTGTGTCTTATAAACTGACTGTTTTTCTTTGCTTGGA
TACTGTGATTCCAAAATAAATCTCATCCAAGCAAGTTAGAGTCCAGCCTAATCAAATGTCATA
ATTGTTGTACCTATTGAAAGTTTTTAAATAATAGATTTATTATGTAAATTATAGTATATGTAA
GTAGCTAATGAAGTAAAGATCATGAAGAAAGAAATTGATAGGTGTAAATGAGAGACCATGTAA
AATATGTAAATTCTAGTACCTGAAATCCTTTCAACAGATTTTTATATAGCAACTGCTCTCTGC
AAGTAGTTAAACTAGAAACTGGGCACATGGTAGAGGCTCACATGGGAGTTGTCCTCACCCTTG
TTAATCTCAAGAAACTCTTATTTATAATAGGTTGCTTCTCTCAGAACTTTTATCTATTACT
TTTTTCTTCTTATGAGTATGTTTACTCTCAGAGTATCTATCTGATGTAGACAGTTGGTGATGC
TTCTGAGACTCAGAATGGTTTACTCTAACAAAACACTGTGCTGTCTATCCCTTGTACTTGCCT
ACTGTAATATGGATTTCACTTCTGAACAGTTTACAGCACAATATTTATTTTAAAGTGAATAAA
ATGTCCACAAGCAAAA
```

FIGURE 94

MEEGGGGVRSLVPGGPVLLVLCGLLEASGGGRALPQLSDDIPFRVNWPGTEFSLPTTGVLYKE
DNYVIMTTAHKEKYKCILPLVTSGDEEEEKDYKGPNPRELLEPLFKQSSCSYRIESYWTYEVC
HGKHIRQYHEEKETGQKININHEYYLGNMLAKNLLFEKEREAEEKEKSNEIPTKNIEGQMTPYY
PVGMGNGTPCSLKQNRPRSSTVMYICHPESKHEILSVAEVTTCEYEVVILTPLLCSHPKYRFR
ASPVNDIFCQSLPGSPFKPLTLRQLEQQEEILRVPFRRNKEGVGWWKYEFCYGKHVHQYHEDK
DSGKTSVVVGTWNQEEHIEWAKKNTARAYHLQDDGTQTVRMVSHFYGNGDICDITDKPRQVTV
KLKCKESDSPHAVTVYMLEPHSCQYILGVESPVICKILDTADENGLLSLPN

Important features of the protein:

Signal peptide:

amino acids 1-30

Glycosaminoglycan attachment site.

amino acids 28-32

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 337-341

N-myristoylation sites.

amino acids 6-12, 23-29, 29-35, 49-55, 141-147, 152-158, 192-198, 196-202

Gram-positive cocci surface proteins 'anchoring' hexapeptide.

amino acids 54-60

FIGURE 95

```
TTCCGTTTCTGGGAGGAGTGAGGGGCAACGGGTCGGAGAAAAAGGAAAAAAGAAGGGCTCAGC
GCCTCCCCGCCGGGCCGTGGACAGAGGGGCACAGTTTCGGCAGGCGGGTGAGGTCGCTGAGGG
CCCGCCGGAGATGTTTTCCTTGTCGAGCACGGTGCAACCCCAGGTTACAGTTCCTCTGAGTCA
TCTCATCAATGCCTTCCATACACCAAAAAACACTTCTGTTTCTCTCAGTGGAGTGTCAGTTTC
TCAAAACCAGCATCGAGATGTAGTTCCTGAGCATGAGGCTCCCAGCAGTGAGCCTTCACTTAA
CTTAAGGGACCTTGGATTATCTGAACTAAAAATTGGACAGATTGATCAGCTGGTAGAAAATCT
ACTTCCTGGATTTTGTAAAGGCAAAAACATTTCTTCCCATTGGCATACATCCCATGTCTCTGC
ACAATCCTTCTTTGAAAATAAATATGGTAACTTAGATATATTTAGTACATTACGTTCCTCTTG
CTTGTATCGACATCATTCAAGAGCTCTTCAAAGCATTTGTTCAGATCTTCAGTACTGGCCAGT
TTTCATACAGTCTCGGGGTTTTAAAACTTTGAAATCAAGGACACGACGTCTCCAGTCTACCTC
CGAGAGATTAGCTGAAACACAGAATATAGCGCCATCATTCGTGAAGGGGTTTCTTTTGCGGGA
CAGAGGATCAGATGTTGAGAGTTTGGACAAACTCATGAAAACCAAAAATATACCTGAAGCTCA
CCAAGATGCATTTAAAACTGGTTTTGCGGAAGGTTTTCTGAAAGCTCAAGCACTCACACAAAA
AACCAATGATTCCCTAAGGCGAACCCGTCTGATTCTCTTCGTTCTGCTGCTATTCGGCATTTA
TGGACTTCTAAAAAACCCATTTTTATCTGTCCGCTTCCGGACAACAACAGGGCTTGATTCTGC
AGTAGATCCTGTCCAGATGAAAAATGTCACCTTTGAACATGTTAAAGGGGTGGAGGAAGCTAA
ACAAGAATTACAGGAAGTTGTTGAATTCTTGAAAAATCCACAAAAATTTACTATTCTTGGAGG
TAAACTTCCAAAAGGAATTCTTTTAGTTGGACCCCCAGGGACTGGAAAGACACTTCTTGCCCG
AGCTGTGGCGGGAGAAGCTGATGTTCCTTTTTATTATGCTTCTGGATCCGAATTTGATGAGAT
GTTTGTGGGTGTGGGAGCCAGCCGTATCAGAAATCTTTTTAGGGAAGCAAAGGCGAATGCTCC
TTGTGTTATATTTATTGATGAATTAGATTCTGTTGGTGGGAAGAGAATTGAATCTCCAATGCA
TCCATATTCAAGGCAGACCATAAATCAACTTCTTGCTGAAATGGATGGTTTTAAACCCAATGA
AGGAGTTATCATAATAGGAGCCACAAACTTCCCAGAGGCATTAGATAATGCCTTAATACGTCC
TGGTCGTTTTGACATGCAAGTTACAGTTCCAAGGCCAGATGTAAAAGGTCGAACAGAAATTTT
GAAATGGTATCTCAATAAAATAAAGTTTGATCAATCCGTTGATCCAGAAATTATAGCTCGAGG
TACTGTTGGCTTTTCCGGAGCAGAGTTGGAGAATCTTGTGAACCAGGCTGCATTAAAAGCAGC
TGTTGATGGAAAAGAAATGGTTACCATGAAGGAGCTGGAGTTTTCCAAAGACAAAATTCTAAT
GGGGCCTGAAAGAAGAAGTGTGGAAATTGATAACAAAAACAAAACCATCACAGCATATCATGA
ATCTGGTCATGCCATTATTGCATATTACACAAAAGATGCAATGCCTATCAACAAAGCTACAAT
CATGCCACGGGGGCCAACACTTGGACATGTGTCCCTGTTACCTGAGAATGACAGATGGAATGA
AACTAGAGCCCAGCTGCTTGCACAAATGGATGTTAGTATGGGAGGAAGAGTGGCAGAGGAGCT
TATATTTGGAACCGACCATATTACAACAGGTGCTTCCAGTGATTTTGATAATGCCACTAAAAT
AGCAAAGCGGATGGTTACCAAATTTGGAATGAGTGAAAAGCTTGGAGTTATGACCTACAGTGA
TACAGGGAAACTAAGTCCAGAAACCCAATCTGCCATCGAACAAGAAATAAGAATCCTTCTAAG
GGACTCATATGAACGAGCAAAACATATCTTGAAAACTCATGCAAAGGAGCATAAGAATCTCGC
AGAAGCTTTATTGACCTATGAGACTTTGGATGCCAAAGAGATTCAAATTGTTCTTGAGGGGAA
AAAGTTGGAAGTGAGATGATAACTCTCTTGATATGGATGCTTGCTGGTTTTATTGCAAGAATA
TAAGTAGCATTGCAGTAGTCTACTTTTACAACGCTTTCCCCTCATTCTTGATGTGGTGTAATT
GAAGGGTGTGAAATGCTTTGTCAATCATTTGTCACATTTATCCAGTTTGGGTTATTCTCATTA
TGACACCTATTGCAAATTAGCATCCCATGGCAAATATATTTTGAAAAAATAAAGAACTATCAG
GATTGAAAACAAAAAAAAAAAA
```

FIGURE 96

MFSLSSTVQPQVTVPLSHLINAFHTPKNTSVSLSGVSVSQNQHRDVVPEHEAPSSEPSLNLRD
LGLSELKIGQIDQLVENLLPGFCKGKNISSHWHTSHVSAQSFFENKYGNLDIFSTLRSSCLYR
HHSRALQSICSDLQYWPVFIQSRGFKTLKSRTRRLQSTSERLAETQNIAPSFVKGFLLRDRGS
DVESLDKLMKTKNIPEAHQDAFKTGFAEGFLKAQALTQKTNDSLRRTRLILFVLLLFGIYGLL
KNPFLSVRFRTTTGLDSAVDPVQMKNVTFEHVKGVEEAKQELQEVVEFLKNPQKFTILGGKLP
KGILLVGPPGTGKTLLARAVAGEADVPFYYASGSEFDEMFVGVGASRIRNLFREAKANAPCVI
FIDELDSVGGKRIESPMHPYSRQTINQLLAEMDGFKPNEGVIIGATNFPEALDNALIRPGRF
DMQVTVPRPDVKGRTEILKWYLNKIKFDQSVDPEIIARGTVGFSGAELENLVNQAALKAAVDG
KEMVTMKELEFSKDKILMGPERRSVEIDNKNKTITAYHESGHAIIAYYTKDAMPINKATIMPR
GPTLGHVSLLPENDRWNETRAQLLAQMDVSMGGRVAEELIFGTDHITTGASSDFDNATKIAKR
MVTKFGMSEKLGVMTYSDTGKLSPETQSAIEQEIRILLRDSYERAKHILKTHAKEHKNLAEAL
LTYETLDAKEIQIVLEGKKLEVR

Important features of the protein:
Transmembrane domain:
amino acids 238-259

N-glycosylation sites.
amino acids 28-32, 90-94, 230-234, 278-282, 535-539, 584-588,
623-627

N-myristoylation sites.
amino acids 35-41, 266-272, 286-292, 325-331, 357-363, 599-605

Amidation site.
amino acids 387-393, 709-713

ATP/GTP-binding site motif A (P-loop).
amino acids 322-330

AAA-protein family proteins
amino acids 315-336, 343-386, 405-451

FIGURE 97

GATGGCGCAGCCACAGCTTCTGTGAGATTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGG

GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAA

ACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT

CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG

ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG

GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAG**ATG**AAAGCCTCTAGT

CTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG

ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT

GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT

GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC

TATCTGGACAGGGTATTTAAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC

AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCATGCCCACATGACA

TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG

GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAG

GAGACAGAA**TAG**GAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT

TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT

GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGATTGTCTTTATGCATCCCC

AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT

TATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTTATTTTTTTACTTGGACATG

AAACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT

ACAGTAAAAAAAAAAAACCTTGTAAATTCTAGAAGAGTGGCTAGGGGGGTTATTCATTTGTAT

TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC

TACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG

ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG

CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAA

FIGURE 98

MKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNGFSEIRGSVQAKDGNIDIR
ILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKISSLANSFLTIKKDLRLC
HAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEETE

Signal sequence:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 107-110, 140-143

N-myristoylation site.

amino acids 51-56

Interleukin 10:

amino acids 9-176

FIGURE 99

GCGCCGGCTCCGCGCCTCGCGCCCAGTCCGCGGGCCGCGCCGCCGCTCCCGCCGCTCCCGCCG

CTCCCGCAGCCGCCCCGCCGCCCGCCCGGAGCCCCGCGTCCCTAGGCCTGGCTCCCGCCTGCC

CGAGACCCGCCCAGCCTGCCCCGCTCAGCCGCCAGAGAAG**ATG**CGGCTGCTCCCGGAATGGTT

CCTCTTGCTCTTTGGCCCGTGGCTCCTTAGGAAGGCCGTCAGTGCCCAGATACCAGAGTCCGG

AAGGCCGCAGTACCTGGGGCTGCGCCCCGCCGCGGCCGGAGCGGGTGCCCCCGGCCAGCAGCT

CCCAGAGCCAAGGTCTTCGGACGGCCTAGGCGTGGGCCGCGCCTGGAGCTGGGCCTGGCCGAC

CAACCACACGGGGGCGCTGGCCCGGGCAGGGGCAGCCGGGGCGTTGCCCGCGCAGCGCACCAA

GAGGAAGCCGTCCATCAAGGCGGCGCGCGCCAAAAGATCTTCGGCTGGGGGGACTTCTACTT

TCGGGTGCATACCCTCAAGTTTTCGCTGCTGGTGACCGGCAAGATCGTGGACCATGTGAACGG

TACCTTCAGTGTGTATTTCCGCCACAACTCGTCCAGCCTGGGCAACCTCAGTGTCAGCATCGT

GCCGCCCTCCAAGCGTGTCGAGTTCGGAGGAGTCTGGCTGCCCGGGCCTGTCCCCCACCCTCT

GCAGTCTACGCTCGCCCTGGAGGGGGTGCTTCCTGGGCTGGGGCCCCGCTGGGGATGGCAGC

AGCAGCGGCGGGGCCGGGGCTTGGGGGCTCCCTCGGGGGCGCACTGGCGGGGCCGCTTGGGGG

CGCGTTGGGAGTGCCTGGGGCCAAAGAGTCACGCGCTTTCAATTGCCACGTGGAGTATGAGAA

GACAAACCGCGCGCGCAAGCACCGACCGTGCCTGTACGACCCGTCGCAGGTGTGTTTCACCGA

GCACACGCAGAGCCAGGCCGCCTGGCTCTGTGCCAAGCCCTTCAAAGTCATCTGTATCTTCGT

CTCTTTCCTCAGCTTTGACTACAAACTGGTGCAGAAGGTGTGCCCAGACTATAACTTCCAGAG

TGAGCACCCCTACTTCGGA**TAG**CGCCCCTCCCCAGCCAGTCCTGAGCCTCCCGCCAAATCCCA

GCCTCACTAGGTGGGACCCCCTTCCCAGTGTTCTGCCGCTCCTGTGGCCATGTCGCCCACTCC

TTCCACTCTGGGGGCGGAGGGGAATGGCTTCTCGGGACCCTCAGCTAGCGTGGGTGCCCTTTT

CCTTATGCGGAGTGCCCGCAAGGCTGGGGTAGCCCCCTCCAGTACACCCCAAAGTGAAAGGGA

TAAGAGTGCAGCCCCAGAATAGGCGGGGCTTGGAGGCGGTCCCAATGTCCCCTGGGTCCACAG

TGGGTCCCCTTTTCACCCTTGGCGCTAGGCTGCGCACTCCCTTTCCCCGCAGCTTTAATAACT

CCTGGCCTGGCACCCTCACCCCACCCTGACTTTCCCATCCCCCAGCGCTTGTCCTGCTTCACC

ATACCCCGCCTAAGACTGTAAAGGCCTAAAAACCTCGGCCTGTCCTCCCACCATTCTGCCTGC

CATATGCCTGTCCCCTTTTCCTCCAAACCCTATTAGGGTACCGGAAGCAGAACCCCTGGGCTG

AGGCCCTGGCCCTGCCCCCGGCCCCTGCCCCTGCCCGCCCCCCTCCAGTCCAGGCAGTCGAGC

TCCACCTGCCCTCTCCTGCTGCTTCCTCTCGGTGATATTTTTTCTACGCCAAAACAGACGGGA

AAGGGAACAAAATAAAGTGAAATCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAA

FIGURE 100

MRLLPEWFLLLFGPWLLRKAVSAQIPESGRPQYLGLRPAAAGAGAPGQQLPEPRSSDGLGVGR
AWSWAWPTNHTGALARAGAAGALPAQRTKRKPSIKAARAKKIFGWGDFYFRVHTLKFSLLVTG
KIVDHVNGTFSVYFRHNSSSLGNLSVSIVPPSKRVEFGGVWLPGPVPHPLQSTLALEGVLPGL
GPPLGMAAAAAGPGLGGSLGGALAGPLGGALGVPGAKESRAFNCHVEYEKTNRARKHRPCLYD
PSQVCFTEHTQSQAAWLCAKPFKVICIFVSFLSFDYKLVQKVCPDYNFQSEHPYFG

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 273-288

N-glycosylation sites.

amino acids 72-76, 133-137, 143-147, 149-153

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 35-41, 58-64, 60-66, 81-87, 84-90, 184-190, 194-200,
203-209, 205-211, 206-212, 209-215, 217-223, 221-227, 224-230

Cytochrome b/b6 Qo site signature.

amino acids 5-11

FIGURE 101

AATGCCCC**ATG**CGCACCCCACAGCTCGCGCTCCTGCAAGTGTTCTTTCTGGTGTTCCCCGATG
GCGTCCGGCCTCAGCCCTCTTCCTCCCCATCAGGGGCAGTGCCCACGTCTTTGGAGCTGCAGC
GAGGGACGGATGGCGGAACCCTCCAGTCCCCTTCAGAGGCGACTGCAACTCGCCCGGCCGTGC
CTGGACTCCCTACAGTGGTCCCTACTCTCGTGACTCCCTCGGCCCCTGGGAATAGGACTGTGG
ACCTCTTCCCAGTCTTACCGATCTGTGTCTGTGACTTGACTCCTGGAGCCTGCGATATAAATT
GCTGCTGCGACAGGGACTGCTATCTTCTCCATCCGAGGACAGTTTTCTCCTTCTGCCTTCCAG
GCAGCGTAAGGTCTTCAAGCTGGGTTTGTGTAGACAACTCTGTTATCTTCAGGAGTAATTCCC
CGTTTCCTTCAAGAGTTTTCATGGATTCTAATGGAATCAGGCAGTTTTGTGTCCATGTGAACA
ACTCAAACTTAAACTATTTCCAGAAGCTTCAAAAGGTCAATGCAACCAACTTCCAGGCCCTGG
CTGCAGAGTTTGGAGGCGAATCATTCACTTCAACATTCCAAACTCAATCACCACCATCTTTTT
ACAGGGCTGGGGACCCCATTCTTACTTACTTCCCCAAGTGGTCTGTAATAAGCTTGCTGAGAC
AACCTGCAGGAGTTGGAGCTGGGGGACTCTGTGCTGAAAGCAATCCTGCAGGTTTCCTAGAGA
GTAAAAGTACAACTTGCACTCGTTTTTTCAAGAACCTGGCTAGTAGCTGTACCTTGGATTCAG
CCCTCAATGCTGCCTCTTACTATAACTTCACAGTCTTAAAGGTTCCAAGAAGCATGACTGATC
CACAGAATATGGAGTTCCAGGTTCCTGTAATACTTACCTCACAGGCTAATGCTCCTCTGTTGG
CTGGAAACACTTGTCAGAATGTAGTTTCTCAGGTCACCTATGAGATAGAGACCAATGGGACTT
TTGGAATCCAGAAAGTTTCTGTCAGTTTGGGACAAACCAACCTGACTGTTGAGCCAGGCGCTT
CCTTACAGCAACACTTCATCCTTCGCTTCAGGGCTTTTCAACAGAGCACAGCTGCTTCTCTCA
CCAGTCCTAGAAGTGGGAATCCTGGCTATATAGTTGGGAAGCCACTCTTGGCTCTGACTGATG
ATATAAGTTACTCAATGACCCTCTTACAGAGCCAGGGTAATGGAAGTTGCTCTGTTAAAAGAC
ATGAAGTGCAGTTTGGAGTGAATGCAATATCTGGATGCAAGCTCAGGTTGAAGAAGGCAGACT
GCAGCCACTTGCAGCAGGAGATTTATCAGACTCTTCATGGAAGGCCCAGACCAGAGTATGTTG
CCATCTTTGGTAATGCTGACCCAGCCCAGAAAGGAGGGTGGACCAGGATCCTCAACAGGCACT
GCAGCATTTCAGCTATAAACTGTACTTCCTGCTGTCTCATACCAGTTTCCCTGGAGATCCAGG
TATTGTGGGCATATGTAGGTCTCCTGTCCAACCCGCAAGCTCATGTATCAGGAGTTCGATTCC
TATACCAGTGCCAGTCTATACAGGATTCTCAGCAAGTTACAGAAGTATCTTTGACAACTCTTG
TGAACTTTGTGGACATTACCCAGAAGCCACAGCCTCCAAGGGGCCAACCCAAAATGGACTGGA
AATGGCCATTCGACTTCTTTCCCTTCAAAGTGGCATTCAGCAGAGGAGTATTCTCTCAAAAAT
GCTCAGTCTCTCCCATCCTTATCCTGTGCCTCTTACTACTTGGAGTTCTCAACCTAGAGACTA
TG**TGA**AGAAAGAAAATAATCAGATTTCAGTTTTCCCTATGAGAAACTCTGAGGCAGCCACTT
ATCTTGGCTAAATAGAACCTCACCTGCTCATGACCAGAGAGCATTTAGGATAATAGATGACCT
AACTGAAGGAATCCTTGTATATGAAGGAGTTATTTTAGAAAAGCAATAAAAATATTTTATTC
ATCNTAAAAAAAAAAA

FIGURE 102

MRTPQLALLQVFFLVFPDGVRPQPSSSPSGAVPTSLELQRGTDGGTLQSPSEATATRPAVPGL

PTVVPTLVTPSAPGNRTVDLFPVLPICVCDLTPGACDINCCCDRDCYLLHPRTVFSFCLPGSV

RSSSWVCVDNSVIFRSNSPFPSRVFMDSNGIRQFCVHVNNSNLNYFQKLQKVNATNFQALAAE

FGGESFTSTFQTQSPPSFYRAGDPILTYFPKWSVISLLRQPAGVGAGGLCAESNPAGFLESKS

TTCTRFFKNLASSCTLDSALNAASYYNFTVLKVPRSMTDPQNMEFQVPVILTSQANAPLLAGN

TCQNVVSQVTYEIETNGTFGIQKVSVSLGQTNLTVEPGASLQQHFILRFRAFQQSTAASLTSP

RSGNPGYIVGKPLLALTDDISYSMTLLQSQGNGSCSVKRHEVQFGVNAISGCKLRLKKADCSH

LQQEIYQTLHGRPRPEYVAIFGNADPAQKGGWTRILNRHCSISAINCTSCCLIPVSLEIQVLW

AYVGLLSNPQAHVSGVRFLYQCQSIQDSQQVTEVSLTTLVNFVDITQKPQPPRGQPKMDWKWP

FDFFPFKVAFSRGVFSQKCSVSPILILCLLLLGVLNLETM

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 484-505, 581-600

N-glycosylation sites.

amino acids 78-82, 165-169, 179-185, 279-285, 331-337, 347-351, 410-414, 487-491

N-myristoylation sites.

amino acids 30-36, 41-47, 124-130, 232-238, 236-242, 409-415

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 420-431

FIGURE 103

CCTAATTCTCAAGGTGATGCTATTTAGGAAGTCATAACTCATGTGAGTGGAGCCATGTGGGAT
TAAGAAGTGATAGGAGAGCTTGCTGTCTGTCTCTGCTCTCCACTGTGTGAGGATACAACAGGA
AGACAGCCATCTGGTGAGGAAGAGAGGGCCCTCGCCAGATACCGGACCTGCTGACACCTTGAT
CTTGGACTTCCCATCTTCCAGGAAGGCCTGACCTCAGTTGTTCCAGGGTAAAGAATTTGGGCA
GTGCCCACACCCACGCTGTTGGATAACATTTCTTCACCATACCAGTGAGGGTGAATGTGTACA
CGCCCAGCTTCCTGCCTGTTACTCTCCACAGT**ATG**CGAAGAATATCCCTGACTTCTAGCCCTG
TGCGCCTTCTTTTGTTTCTGCTGTTGCTACTAATAGCCTTGGAGATCATGGTTGGTGGTCACT
CTCTTTGCTTCAACTTCACTATAAATCATTGTCCAGACCTGGACAGCCCTGGTGTGAAGCGC
AGGTCTTCTTGAATAAAAATCTTTTCCTTCAGTACAACAGTGACAACAACATGGTCAAACCTC
TGGGCCTCCTGGGGAAGAAGGTATATGCCACCAGCACTTGGGGAGAATTGACCCAAACGCTGG
GAGAAGTGGGGCGAGACCTCAGGATGCTCCTTTGTGACATCAAACCCCAGATAAAGACCAGTG
ATCCTTCCACTCTGCAAGTCGAGATGTTTTGTCAACGTGAAGCAGAACGGTGCACTGGTGCAT
CCTGGCAGTTCGCCACCAATGGAGAGAAATCCCTCCTCTTTGACGCAATGAACATGACCTGGA
CAGTAATTAATCATGAAGCCAGTAAGATCAAGGAGACATGGAAGAAAGACAGAGGGCTGGAAA
AGTATTTCAGGAAGCTCTCAAAGGGAGACTGCGATCACTGGCTCAGGGAATTCTTAGGGCACT
GGGAGGCAATGCCAGAACCGACAGGCAGAAGATCCACC**TAG**AGGTGATACCACGGCGGCGCAG
AGTTGTTCACCTGTGGTCCTCGATCGCTGACAGCCTTGGCTCCCACTGCTGTGTGTTCCCTGA
GTCAAGTGGAGGCGGAGCCTGCAATGAGCGGAGATCGCGCCTCTGCATTCCAGTCTTGGCAAC
AGAGCAAGACTCCGTCTCAAAAAAAAAAAATTTTTTTTCAGTACATATTTTTTAAAAGATAGG
GCTGGGCACAGCAGCTCACATCTATAATCCCAACACTTTGGGAGGCCTAGGCAGGAGGATCAC
TTGAGCCCAGGAATCTGAAGCTGCAGTGAGCCTTTGCTCGTGAGATTGTGGACCTATGATCCT
ACCACCAGCCCACCTGGTTCTAACACCCCCTCCTCTATGTGTGAGAGGGAGAGAAGAAAAGTG
AGGGAGAAAAGAGAGATAAGCAAAGAACAGAGAGGAAAAATGGAAAATAAGAGGAAATTGGGG
GAATTAAACAGAGGGGAGGGCATGGATCCCCGGGAGTTAGAAGAGTAGCAGCTTGTGGATTAC
TACGCAGTGGAGGAAGAAGAGTTGTTGGAAATTATTTGAGAGGTAGTATAATCATTTGTGAGG
CAGTTTTCTGCATTCACCATTTCTCACAGACTAAGTTACTCATAAGCAAACGTGCAATTCACA
TTACACTGAAATTCTTCCCTAATACATCATTTGCATTGGAATAAAGTACGGTTTTCAAACAAC
CTGATATAGCAGAACTGACTGTATAAATTATGTGAGCACAGTGCAAGTAATTCTTTGTTTGTT
TGTTTGTTTTTTTGAGACAGAGTCTCACTCTATCTCCCAGGCTGGAGTGTAGTGGTGCGATCC
CGGCTCACTGCAACCTCGATCTCCCAGGCTCAAGCGATTCCCCTGCCTCAGCCTCCTGAGTAG
CTGGGATTACAGGCATGAGCCACCACGCCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGT
TTCACCCTGTTGGCCAGGCTGGTCTCGAACTACGGACCTCAGGTGATCTGCCCCCCTCAGCCT
CTCAAAGTGCTGGGATTATAGCATGAGCCACTGAGCCCAGACACAAGTAGTTCTTTCTGATAA
ACACTTTAACACTGAATGCA

FIGURE 104

MRRISLTSSPVRLLLFLLLLLIALEIMVGGHSLCFNFTIKSLSRPGQPWCEAQVFLNKNLFLQ

YNSDNNMVKPLGLLGKKVYATSTWGELTQTLGEVGRDLRMLLCDIKPQIKTSDPSTLQVEMFC

QREAERCTGASWQFATNGEKSLLFDAMNMTWTVINHEASKIKETWKKDRGLEKYFRKLSKGDC

DHWLREFLGHWEAMPEPTGRRST

Important features of the protein:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 11-30(possible type II protein)

N-glycosylation site.

amino acids 36-39, 154-157

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 2-5, 182-185, 209-212

Casein kinase II phosphorylation site.

amino acids 86-89, 93-96, 142-145, 185-188

N-myristoylation site.

amino acids 46-51

Amidation site.

amino acids 77-80, 207-210

FIGURE 105

TTTTCCGAGTGACCTTCTTG**ATG**CTGGCTGTTTCTCTCACCGTTCCCCTGCTTGGAGCCATGA

TGCTGCTGGAATCTCCTATAGATCCACAGCCTCTCAGCTTCAAAGAACCCCCGCTCTTGCTTG

GTGTTCTGCATCCAAATACGAAGCTGCGACAGGCAGAAAGGCTGTTTGAAAATCAACTTGTTG

GACCGGAGTCCATAGCACATATTGGGGATGTGATGTTTACTGGGACAGCAGATGGCCGGGTCG

TAAAACTTGAAAATGGTGAAATAGAGACCATTGCCCGGTTTGGTTCGGGCCCTTGCAAAACCC

GAGATGATGAGCCTGTGTGTGGGAGACCCCTGGGTATCCGTGCAGGGCCCAATGGGACTCTCT

TTGTGGCCGATGCATACAAGGGACTATTTGAAGTAAATCCCTGGAAACGTGAAGTGAAACTGC

TGCTGTCCTCCGAGACACCCATTGAGGGGAAGAACATGTCCTTTGTGAATGATCTTACAGTCA

CTCAGGATGGGAGGAAGATTTATTTCACCGATTCTAGCAGCAAATGGCAAAGACGAGACTACC

TGCTTCTGGTGATGGAGGGCACAGATGACGGGCGCCTGCTGGAGTATGATACTGTGACCAGGG

AAGTAAAAGTTTTATTGGACCAGCTGCGGTTCCCGAATGGAGTCCAGCTGTCTCCTGCAGAAG

ACTTTGTCCTGGTGGCAGAAACAACCATGGCCAGGATACGAAGAGTCTACGTTTCTGGCCTGA

TGAAGGGCGGGGCTGATCTGTTTGTGGAGAACATGCCTGGATTTCCAGACAACATCCGGCCCA

GCAGCTCTGGGGGGTACTGGGTGGGCATGTCGACCATCCGCCCTAACCCTGGGTTTTCCATGC

TGGATTTCTTATCTGAGAGACCCTGGATTAAAAGGATGATTTTTAAGCTCTTTAGTCAAGAGA

CGGTGATGAAGTTTGTGCCGCGGTACAGCCTCGTCCTAGAACTCAGCGACAGCGGTGCCTTCC

GGAGAAGCCTGCATGATCCCGATGGGCTGGTGGCCACCTACATCAGCGAGGTGCACGAACACG

ATGGGCACCTGTACCTGGGCTCTTTCAGGTCCCCCTTCCTCTGCAGACTCAGCCTCCAGGCTG

TT**TAG**CCCTCCCAGATAGCTGCCCCTGCCACGCAGGCCAGGAGTCTTCACACTCAGGCACCAG

GCCTGGTCCAGGAGGAGCTGTGGACACAGTCGTGGTTCAAGTGTCCACATGCACCTGTTAGTC

CCTGAGAGGTGGTGGGAATGGCTGCTTCATTCCTCGAGGATGCCCGGGCCCCACCTGGGCTTG

TCTTTCTGTTTAGAGGGAAGTGTAACATATCTGCCATGAGGAACATAAATTCATGTAAAGCCA

TTTTCTCTTAAACAAAACAAAACTTTCTAAGTACAATCATTCTCTAGGATTTGGGAAGCTCCT

TGCACTTGGAACAGGGCTCAGGTGGGTGGAGCAGTAAGGCACTACCCAGAGAGCTTGCTGCTG

CGGCCCTGTCCTGCGGCCTCAAAGTTCTTCTTTACTATATATAACGTGCGGTCATACCTTTCT

TCGTTGTGGTGGGGATGGAAGAGCAGAGGGAGCATGGCCCAGGGGTGTTGAGGCCAGCGGTGA

GAGCCGTGTTAGCCAAGACATGGAACTGTGTTCTCAAGGGTTATGTGGGCGTGGGCTCTCCA

TAGTGTGTATGAAAAGCTTGTTGACTCTAGCGGCTCAGAGAGGACTTTGCTGGGTTTCTTTCT

GTGAATATCTCCGTGCTGACCATGCTGGAATTGGATGATTCTGCAATTCGGGACCTACTGCAG

GGGTCCGTTTAGTAACGTCTTGTCTGTGATCTTTGTTCTTGACCTCTAGACCCCAAGATGTGA

ACAGTGCACGTGTTAATGTCATCTTTGCTCATGTGTTATAAGCCCCAAGTTGCTGTATATTTT

CACAAGTATGTCTACACACTGG

FIGURE 106

MLAVSLTVPLLGAMMLLESPIDPQPLSFKEPPLLLGVLHPNTKLRQAERLFENQLVGPESIAH
IGDVMFTGTADGRVVKLENGEIETIARFGSGPCKTRDDEPVCGRPLGIRAGPNGTLFVADAYK
GLFEVNPWKREVKLLLSSETPIEGKNMSFVNDLTVTQDGRKIYFTDSSSKWQRRDYLLLVMEG
TDDGRLLEYDTVTREVKVLLDQLRFPNGVQLSPAEDFVLVAETTMARIRRVYVSGLMKGGADL
FVENMPGFPDNIRPSSSGGYWVGMSTIRPNPGFSMLDFLSERPWIKRMIFKLFSQETVMKFVP
RYSLVLELSDSGAFRRSLHDPDGLVATYISEVHEHDGHLYLGSFRSPFLCRLSLQAV
                                                   .

Important features of the protein:
Signal peptide:
amino acids 1-13

Transmembrane domain:
amino acids 1-21 (possible type II)

N-glycosylation sites.
amino acids 116-119, 152-155

Casein kinase II phosphorylation sites.
amino acids 19-22, 27-30, 98-101, 146-149, 221-224, 286-289, 332-335

N-myristoylation sites.
amino acids 71-76, 92-97, 189-194, 244-249, 338-343

Amidation site.
amino acids 164-167

FIGURE 107

AACGAAGCGTGCGCGCTTTGGTAACCGGCTAGAAATCCCGCACGCGCGCCTGCCTCCTCTCCC
CAGGCCTGAGCTGCCCCTCCCACTGCCTTTCCTTCTTCCCGCGAGTCAGAAGCTTCGCGAGGG
CCCAGAGAGGCGGTGGGGTGGGCGACCCTACGCCAGCTCCGGGCGGGAGAAAGCCCACCCTCT
CCCGCGCCCCAGGAAACCGCCGGCGTTCGGCGCTGCGCAGAGCC**ATG**GAATTCTCCTGGCTGG
AGACGCGCTGGGCGCGGCCCTTTTACCTGGCGTTCGTGTTCTGCCTGGCCCTGGGGCTGCTGC
AGGCCATTAAGCTGTACCTGCGGAGGCAGCGGCTGCTGCGGGACCTGCGCCCCTTCCCAGCGC
CCCCCACCCACTGGTTCCTTGGGCACCAGAAGTTTATTCAGGATGATAACATGGAGAAGCTTG
AGGAAATTATTGAAAAATACCCTCGTGCCTTCCCTTTCTGGATTGGGCCCTTTCAGGCATTTT
TCTGTATCTATGACCCAGACTATGCAAAGACACTTCTGAGCAGAACAGATCCCAAGTCCCAGT
ACCTGCAGAAATTCTCACCTCCACTTCTTGGAAAAGGACTAGCGGCTCTAGACGGACCCAAGT
GGTTCCAGCATCGTCGCCTACTAACTCCTGGATTCCATTTTAACATCCTGAAAGCATACATTG
AGGTGATGGCTCATTCTGTGAAAATGATGCTGGATAAGTGGGAGAAGATTTGCAGCACTCAGG
ACACAAGCGTGGAGGTCTATGAGCACATCAACTCGATGTCTCTGGATATAATCATGAAATGCG
CTTTCAGCAAGGAGACCAACTGCCAGACAAACAGCACCCATGATCCTTATGCAAAGCCATAT
TTGAACTCAGCAAATCATATTTCACCGCTTGTACAGTTTGTTGTATCACAGTGACATAATTT
TCAAACTCAGCCCTCAGGGCTACCGCTTCCAGAAGTTAAGCCGAGTGTTGAATCAGTACACAG
ATACAATAATCCAGGAAAGAAAGAAATCCCTCCAGGCTGGGGTAAAGCAGGATAACACTCCGA
AGAGGAAGTACCAGGATTTTCTGGATATTGTCCTTTCTGCCAAGGATGAAAGTGGTAGCAGCT
TCTCAGATATTGATGTACACTCTGAAGTGAGCACATTCCTGTTGGCAGGACATGACACCTTGG
CAGCAAGCATCTCCTGGATCCTTTACTGCCTGGCTCTGAACCCTGAGCATCAAGAGAGATGCC
GGGAGGAGGTCAGGGGCATCCTGGGGGATGGGTCTTCTATCACTTGGGACCAGCTGGGTGAGA
TGTCGTACACCACAATGTGCATCAAGGAGACGTGCCGATTGATTCCTGCAGTCCCGTCCATTT
CCAGAGATCTCAGCAAGCCACTTACCTTCCCAGATGGATGCACATTGCCTGCAGGGATCACCG
TGGTTCTTAGTATTTGGGGTCTTCACCACAACCCTGCTGTCTGGAAAAACCCAAAGGTCTTTG
ACCCCTTGAGGTTCTCTCAGGAGAATTCTGATCAGAGACACCCCTATGCCTACTTACCATTCT
CAGCTGGATCAAGGAACTGCATTGGGCAGGAGTTTGCCATGATTGAGTTAAAGGTAACCATTG
CCTTGATTCTGCTCCACTTCAGAGTGACTCCAGACCCCACCAGGCCTCTTACTTTCCCCAACC
ATTTTATCCTCAAGCCCAAGAATGGGATGTATTTGCACCTGAAGAAACTCTCTGAATGT**TAG**A
TCTCAGGGTACAATGATTAAACGTACTTTGTTTTTCGAAGTTAAATTTACAGCTAATGATCCA
AGCAGATAGAAAGGGATCAATGTATGGTGGGAGGATTGGAGGTTGGTGGGATAGGGGTCTCTG
TGAAGAGATCCAAAATCATTTCTAGGTACACAGTGTGTCAGCTAGATCTGTTTCTATATAACT
TTGGGAGATTTTCAGATCTTTTCTGTTAAACTTTCACTACTATTAATGCTGTATACACCAATA
GACTTTCATATATTTTCTGTTGTTTTTAAAATAGTTTTCAGAATTATGCAAGTAATAAGTGCA
TGTATGCTCACTGTCAAAAATTCCCAACACTAGAAATCATGTAGAATAAAAATTTTAAATCT
CACTTCACTTAGCCGACATTCCATGCCCTGACCAATCCTACTGCTTTTCCTAAAAACAGAATA
ATTTGGTGTGCATTCTTTCAGACTTTTTCCTATACATTTTATATGTAGAAATGTAGCAATGTA
TTTGTATAGATGTGATCATTCCTATATTGTTATTGATTTTTTTCACTTAATAAAAATTCACCT
TATTCCTTAAAA

FIGURE 108

MEFSWLETRWARPFYLAFVFCLALGLLQAIKLYLRRQRLLRDLRPFPAPPTHWFLGHQKFIQD
DNMEKLEEIIEKYPRAFPFWIGPFQAFFCIYDPDYAKTLLSRTDPKSQYLQKFSPPLLGKGLA
ALDGPKWFQHRRLLTPGFHFNILKAYIEVMAHSVKMMLDKWEKICSTQDTSVEVYEHINSMSL
DIIMKCAFSKETNCQTNSTHDPYAKAIFELSKIIFHRLYSLLYHSDIIFKLSPQGYRFQKLSR
VLNQYTDTIIQERKKSLQAGVKQDNTPKRKYQDFLDIVLSAKDESGSSFSDIDVHSEVSTFLL
AGHDTLAASISWILYCLALNPEHQERCREEVRGILGDGSSITWDQLGEMSYTTMCIKETCRLI
PAVPSISRDLSKPLTFPDGCTLPAGITVVLSIWGLHHNPAVWKNPKVFDPLRFSQENSDQRHP
YAYLPFSAGSRNCIGQEFAMIELKVTIALILLHFRVTPDPTRPLTFPNHFILKPKNGMYLHLK
KLSEC

Important features of the protein:
Signal peptide:
amino acids 1-29

Transmembrane domains:
amino acids 310-330, 397-413, 459-473

N-glycosylation site.
amino acids 206-210

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 265-269, 504-520

N-myristoylation sites.
amino acids 25-31, 298-304, 353-359, 450-456, 456-462

Cytochrome P450 cysteine heme-iron ligand signature.
amino acids 447-457

Cytochrome P450 cysteine heme-iron ligand proteins.
amino acids 444-475

FIGURE 109

GGCGTTCCGGGCCTCAACTTTGGCGTCGTGAGATTCTTGTGAGGCGTCTGCCTGGAAGCCGGC

AGCAATTTTGCTTCTTTAAAGAGAAAAAGAAGGCTAGGGACTCAGATTCCTGGATTCTGAGAT

CCAGACCAGCTCCTCCCAGACCTCTCCAGAAGAAGCC**ATG**GGAACCCCTCGTATCCAGCATTT

GCTGATCCTCCTGGTCCTAGGAGCCTCCCTCCTGACCTCGGGCCTAGAGCTGTATTGTCAAAA

GGGTCTGTCCATGACTGTGGAAGCAGATCCAGCCAATATGTTTAACTGGACCACAGAGGAAGT

GGAGACTTGTGACAAAGGGGCACTTTGCCAGGAAACCATACTAATAATTAAAGCAGGGACTGA

GACAGCCATTTTGGCCACGAAGGGCTGCATCCCGGAAGGGGAGGAGGCCATAACAATTGTCCA

GCACTCTTCACCTCCCGGCCTGATCGTGACCTCCTACAGTAACTACTGTGAGGATTCCTTCTG

TAATGACAAAGACAGCCTGTCTCAGTTTTGGGAGTTCAGTGAGACCACAGCTTCCACTGTGTC

AACAACCCTCCATTGTCCAACCTGTGTGGCTTTGGGGACCTGTTTCAGTGCTCCTTCTCTTCC

CTGTCCCAATGGTACAACTCGATGCTATCAAGGAAAACTTGAGATCACTGGAGGTGGCATTGA

GTCGTCTGTGGAGGTCAAAGGCTGTACAGCCATGATTGGCTGCAGGCTGATGTCTGGAATCTT

AGCAGTAGGACCCATGTTTGTGAGGGAAGCGTGCCCACATCAGCTGCTCACTCAACCTCGAAA

GACTGAAAATGGGGCCACCTGTCTTCCCATTCCTGTTTGGGGGTTACAGCTACTGCTGCCATT

GCTGCTGCCATCATTTATTCACTTTTCC**TAA**GAAGGCACTTCTGGGCCTGGGTCTGAGGACAT

CTTTTTTGACTGGGAGCCTTCTTACTGTTGAGGTTCAACAAGCTGAGGAGTAGATGGGAATTT

GAGGGAGAATACAGAGATACTATGAACGTATTTGACATTTTTAATACAATTTCTGCTATAATT

TTTGTATGCAGTAGGCGTTACTAATAAACATTTCTGCTGTGA

FIGURE 110

MGTPRIQHLLILLVLGASLLTSGLELYCQKGLSMTVEADPANMFNWTTEEVETCDKGALCQET

ILIIKAGTETAILATKGCIPEGEEAITIVQHSSPPGLIVTSYSNYCEDSFCNDKDSLSQFWEF

SETTASTVSTTLHCPTCVALGTCFSAPSLPCPNGTTRCYQGKLEITGGGIESSVEVKGCTAMI

GCRLMSGILAVGPMFVREACPHQLLTQPRKTENGATCLPIPVWGLQLLLPLLLPSFIHFS

Important features of the protein:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 184-201

N-glycosylation sites.

amino acids 45-49, 159-163

N-myristoylation sites.

amino acids 31-37, 70-76, 99-105, 147-153, 160-166, 174-180, 175-181

FIGURE 111

```
CGAGAAGAGGACAGAGGAGACTGAGCAAAGGGGGGTGGGCTCCAGGCGACCCCTAGCCCAATTCTGCCCCTCCAT
CCCAAGGGGCAGAGAAATTGTCTTTCTTTGCTGACTCCTACGAGGAAAAAAAAAAAAAAAAAAAAAAACCATTTAA
AGGGAAAGATAAACGGAGACGGAGGAAAGGTGGCAGCCAGATTACTTAGAGAGGCACAGAGGAGAGAGATCGGGG
TGAGTCGCCATGGGGACTCCCAGGGCCCAGCACCCGCCGCCTCCCCAGCTGCTGTTCCTAATTCTGCTGAGCTGT
CCCTGGATCCAGGGTCTGCCCCTGAAGGAGGAGGAGATATTGCCAGAGCCTGGAAGTGAGACCCCCACGGTGGCC
TCTGAGGCCCTGGCTGAACTGCTTCATGGGGCCCTGCTGAGGAGGGGCCCAGAGATGGGCTACCTGCCAGGATCT
GATCCGGACCCCACGCTAGCCACCCCTCCGGCCGGCCAGACTCTCGCAGTGCCCTCCCTGCCACGGGCCACTGAG
CCGGGGACAGGGCCTCTGACAACAGCCGTCACCCCTAACGGGGTCAGGGGGGCAGGCCCCACTGCGCCAGAACTG
CTGACCCCGCCCCCAGGAACCACAGCCCCACCCCCACCCAGCCCTGCCTCCCCAGGGCCTCCCCTTGGGCCTGAG
GGAGGAGAGGAGGAGACGACGACCACCATCATCACCACGACAACTGTTACCACTACGGTGACCAGCCCAGTTCTG
TGTAATAACAACATCTCCGAGGGCGAAGGGTATGTGGAGTCTCCAGATCTGGGGAGCCCCGTCAGCCGCACCCTG
GGGCTCCTGGACTGCACTTACAGCATCCATGTCTACCCTGGCTACGGCATTGAGATCCAGGTGCAGACGCTGAAC
CTGTCACAGGAAGAGGAGCTCCTGGTGCTGGCTGGTGGGGGATCCCCAGGCCTGGCCCCCCGACTCCTGGCCAAC
TCATCCATGCTTGGAGAAGGACAAGTCCTTCGGAGCCCAACCAACCGGCTGCTTCTGCACTTCCAGAGCCCACGG
GTCCCAAGGGGCGGTGGCTTCAGGATCCACTATCAGGCCTACCTCCTGAGCTGTGGCTTCCCTCCCCGGCCGGCC
CATGGGGACGTGAGTGTGACGGACCTGCACCCTGGGGGCACTGCCACCTTTCACTGTGATTCGGGCTACCAGCTG
CAGGGAGAGGAGACCCTCATCTGCCTCAATGGCACCCGGCCATCCTGGAACGGTGAAACCCCCAGCTGCATGGCA
TCCTGTGGTGGCACCATCCACAATGCCACCCTGGGCCGCATCGTGTCCCCAGAGCCTGGGGGGAGCCGTAGGGCCC
AACCTCACCTGCCGTTGGGTCATTGAAGCAGCTGAGGGGCGCCGGCTGCACCTGCACTTTGAAAGGGTCTCGCTG
GATGAGGACAATGACCGGCTGATGGTGCGCTCAGGGGGCAGCCCCCTATCCCCCGTGATCTATGATTCGGACATG
GACGATGTCCCCGAGCGGGGTCTCATCAGTGACGCCCAGTCCCTCTACGTGGAGCTGCTGTCAGAGACACCTGCC
AATCCCCTGCTGTTAAGCCTTCGATTTGAAGCCTTTGAGGAGGATCGCTGCTTCGCCCCCTTCCTGGCACATGGA
AATGTCACTACCACGGACCCTGAGTATCGCCCAGGGGCACTGGCAACCTTCTCGTGCCTCCCAGGATATGCCCTG
GAGCCCCCTGGGCCCCCCAATGCCATCGAATGTGTGGATCCCACAGAACCCCACTGGAACGACACAGAGCCGGCC
TGCAAAGCCATGTGTGGAGGGGAGCTGTCGGAACCAGCTGGCGTGGTCCTCTCTCCCGACTGGCCCCAGAGCTAT
AGCCCGGGCCAAGACTGCGTGTGGGGCGTGCACGTCCAGGAAGAGAAGCGCATCTTGCTCCAAGTTGAGATATTG
AATGTGCGGGAAGGGGACATGCTGACGCTGTTCGACGGGGACGGTCCCAGCGCCCGAGTCTTGGCCCAGCTGCGG
GGACCTCAGCCGCGCCGCCGCCTTCTCTCCTCTGGGCCCGACCTCACACTGCAGTTTCAGGCACCGCCCGGGCCC
CCAAATCCAGGCCTGGGCCAGGGCTTCGTATTGCACTTCAAAGAGGTCCCGAGGAACGACACGTGCCCCGAGCTG
CCACCTCCGGAGTGGGGCTGGAGAACGGCATCCCACGGGGACCTGATCCGGGGCACGGTGCTCACCTACCAGTGC
GAGCCTGGCTACGAGCTGCTAGGCTCCGACATTCTCACTTGCCAGTGGGACCTGTCTTGGAGCGCCGCGCCGCCC
GCCTGCCAAAAGATCATGACTTGTGCTGACCCTGGCGAGATTGCCAACGGGCACCGCACCGCCTCGGACGCCGGC
TTCCCCGTTGGCTCCCACGTCCAGTACCGCTGCCTGCCCAGGGTACAGCCTCGAGGGGGGCAGCCATGCTCACCTGC
TACAGCCGGGACACAGGCACACCCAAGTGGAGCGATAGGGTCCCCAAATGCCCTTGAAGTACGAGCCGTGCCTG
AACCCGGGGGTTCCCGAGAATGGCTACCAGACGCTGTACAAGCACCACTACCAGGCGGGCGAGTCTCTGCGCTTC
TTCTGCTATGAGGGCTTTGAGCTTATCGGCGAGGTCACCATCACCTGTGTGCCCGGCCACCCCTCCCAGTGGACC
AGCCAGCCCCCACTCTGCAAAGTGACCCAGACCACAGATCCATCACGGCAGCTGGAAGGGGGGAACCTGGCCCTG
GCCATCCTGCTGCCTCTAGGCTTGGTCATTGTCCTCGGCAGTGGCGTTTACATCTACTACACCAAGCTTCAGGGA
AAGTCCCTTTTCGGCTTCTCGGGCTCCCACTCCTACAGCCCCATCACCGTGGAGTCGGACTTCAGCAACCCGCTG
TATGAAGCTGGGGATACGCGGGAGTATGAAGTTTCCATCTGAACCCCAAGACTACAGCTGCAGGACCCAGGACGC
CCCTCCCCTCCTCATTCGGGCAGAGGGAAATACGGGACCCGGTCTCTGCCTCCTGGCTGCCCTCCTCCCTGGCTG
TGTAAATAGTCTCCCTATCCCACGAGGGGGCTTTGATGGCCCTGGAGATCCTACAGTAAATAAACCAGCATCCTG
CCGCCCAAAAAA
```

FIGURE 112

```
MGTPRAQHPPPPQLLFLILLSCPWIQGLPLKEEEILPEPGSETPTVASEALAELLHGALLRRG
PEMGYLPGSDPDPTLATPPAGQTLAVPSLPRATEPGTGPLTTAVTPNGVRGAGPTAPELLTPP
PGTTAPPPPSPASPGPPLGPEGGEEETTTTIITTTTVTTTVTSPVLCNNNISEGEGYVESPDL
GSPVSRTLGLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELLVLAGGGSPGLAPRLLANSSMLG
EGQVLRSPTNRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFPPRPAHGDVSVTDLHPGGTATFH
CDSGYQLQGEETLICLNGTRPSWNGETPSCMASCGGTIHNATLGRIVSPEPGGAVGPNLTCRW
VIEAAEGRRLHLHFERVSLDEDNDRLMVRSGGSPLSPVIYDSDMDDVPERGLISDAQSLYVEL
LSETPANPLLLSLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPGALATFSCLPGYALEPPGPPN
AIECVDPTEPHWNDTEPACKAMCGGELSEPAGVVLSPDWPQSYSPGQDCVWGVHVQEEKRILL
QVEILNVREGDMLTLFDGDGPSARVLAQLRGPQPRRRLLSSGPDLTLQFQAPPGPPNPGLGQG
FVLHFKEVPRNDTCPELPPPEWGWRTASHGDLIRGTVLTYQCEPGYELLGSDILTCQWDLSWS
AAPPACQKIMTCADPGEIANGHRTASDAGFPVGSHVQYRCLPGYSLEGAAMLTCYSRDTGTPK
WSDRVPKCALKYEPCLNPGVPENGYQTLYKHHYQAGESLRFFCYEGFELIGEVTITCVPGHPS
QWTSQPPLCKVTQTTDPSRQLEGGNLALAILLPLGLVIVLGSGVYIYYTKLQGKSLFGFSGSH
SYSPITVESDFSNPLYEAGDTREYEVSI
```

Important features of the protein:

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 842-864

N-glycosylation sites.

amino acids 176-180, 222-226, 247-251, 332-336, 355-359, 373-377, 473-477, 517-521, 641-645

Tyrosine kinase phosphorylation site.

amino acids 61-69

N-myristoylation sites.

amino acids 2-8, 84-90, 111-117, 114-120, 190-196, 198-204, 235-241, 309-315, 333-339, 351-357, 472-478, 484-490, 528-534, 626-632, 665-671, 775-781, 842-848

Amidation site.

amino acids 384-388

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23

CUB domain proteins profile.

amino acids 202-218, 376-392, 553-569

FIGURE 113

```
GCCGCGGGCGGAGCTGCCTGCCGGTCCCGCGCCGCGCGTCCGCACTCCTCGGCCCTCGGGCGGTCGATGGGACGG
GGCGCCGCGGAGCAGGAGGCGGCGCCCGTCGGGGTGCTCGGGCCGCGCGGGAGCCCACTGTGGGGCTCGGGCATG
GCGGGCCGCAGGACCTGAGCTCTCCTCAGGGGAGCGGGGAGGCAGCTGCTGGCCGGCGATGGGGACGGAGTGGGG
CCGTCGCCGCCGCGCCGAGCCGTGAGCGCCGAGCCACCGCCGCCGCTACCTCAGCCCTTCGCGAAGCGCCGGGCA
GCTCGGGAAC**ATG**GCCCTGGAGCGGCTCTGCTCGGTCCTCAAAGTGTTGTTAATAACAGTACTGGTAGTGGAAGG
GATTGCCGTGGCCCAAAAAACCCAAGATGGACAAAATATTGGAATCAAGCATATTCCTGCAACCCAGTGTGGCAT
TTGGGTTCGAACCAGCAATGGAGGTCATTTTGCTTCGCCAAATTATCCTGACTCATATCCACCAAACAAGGAGTG
TATCTACATTTTGGAAGCTGCTCCACGTCAAAGAATAGAGTTGACCTTTGATGAACATTATTATATAGAACCATC
ATTTGAGTGTCGGTTTGATCACTTGGAAGTTCGAGATGGGCCATTTGGTTTCTCTCCTCTTATAGATCGTTACTG
TGGCGTGAAAAGCCCTCCATTAATTAGATCAACAGGGAGATTCATGTGGATTAAGTTTAGTTCTGATGAAGAGCT
TGAAGGACTGGGATTTCGAGCAAAATATTCATTTATTCCAGATCCAGACTTTACTTACCTAGGAGGTATTTTAAA
TCCCATTCCAGATTGTCAGTTCGAGCTCTCGGGAGCTGATGGAATAGTGCGCTCTAGTCAGGTAGAACAAGAGGA
GAAAACAAAACCAGGCCAAGCCGTTGATTGCATCTGGACCATTAAAGCCACTCCAAAAGCTAAGATTTATTTGAG
GTTCCTAGATTATCAAATGGAGCACTCAAATGAATGCAAGAGAAACTTCGTTGCAGTCTATGATGGAAGCAGTTC
TATTGAAAACCTGAAGGCCAAGTTTTGCAGCACTGTGGCCAATGATGTAATGCTTAAAACAGGAATTGGAGTGAT
TCGAATGTGGGCAGATGAAGGTAGTCGGCTTAGCAGGTTTCGAATGCTCTTTACTTCCTTTGTGGAGCCTCCCTG
CACAAGCAGCACTTTCTTTTGCCATAGCAACATGTGCATCAATAATTCTTTAGTCTGTAATGGTGTCCAAAATTG
TGCATACCCTTGGGATGAAAATCATTGTAAAGAAAAGAAAAAAGCAGGAGTATTTGAACAAATCACTAAGACTCA
TGGAACAATTATTGGCATTACTTCAGGGATTGTCTTGGTCCTTCTCATTATTTCTATTTTAGTACAAGTGAAACA
GCCTCGAAAAAAGGTCATGGCTTGCAAAACCGCTTTTAATAAAACCGGGTTCCAAGAAGTGTTTGATCCTCCTCA
TTATGAACTGTTTTCACTAAGGGACAAAGAGATTTCTGCAGACCTGGCAGACTTGTCGGAAGAATTGGACAACTA
CCAGAAGATGCGGCGCTCCTCCACCGCCTCCCGCTGCATCCACGACCACCACTGTGGGTCGCAGGCCTCCAGCGT
CAAACAAAGCAGGACCAACCTCAGTTCCATGGAACTTCCTTTCCGAAATGACTTTGCACAACCACAGCCAATGAA
AACATTTAATAGCACCTTCAAGAAAAGTAGTTACACTTTCAAACAGGGACATGAGTGCCCTGAGCAGGCCCTGGA
AGACCGAGTAATGGAGGAGATTCCCTGTGAAATTTATGTCAGGGGGCGAGAAGATTCTGCACAAGCATCCATATC
CATTGACTTC**TAA**TCTTCTGCTAATGGTGATGTGAATTCTTAGGGTGTGTACGTACGCAGCCTCCAGGGCACCAT
ACTGTTTCCAGCAGCCAACCCTTTTCTCCCATCACAACTACGAAGACCTTGATTTACCGTTAACCTATTGTATGG
TGATGTTTTTATTCTCTCAGGCAGTCTATATATGTTAAACCAATCAAGGAACTTACTCTATTCAGTGGAAACAAT
AATCATCTCTATTGCTTGGTGTCATTTATAGGAAGCACTGCCAGTTAAAGAGCATTAGAAGAGGTGGTTGGATGG
AGCCAGGCTCAGGCTGCCTCTTCGTTTTAGCAACAAGAAGACTGCTCTTGACTGATAACAGCTCTGTCAATATTT
TGATGCCACAATAAACTTGATTTTTTTTACATTCCTTTTATTTTTCCTTTCTCTAAATTTAATTTGTTTTATAA
GCCTATCGTTTTACCATTTCATTTTCTTACATAAGTACAAGTGGTTAATGTACCACATACTTCAGTATAGGCATT
TGTTCTTGAGTGTGTCAAAATACAGCTAGTTACTGTGCCAATTAAGACCCAGTTGTATTTCACCCATCTGTTTCT
TCTTGGCTAATCTCTGTACTTCTGCCTTTTAATTACTGGGCCCTTATTCCTTATTTTCTGTGAGAAATAATAGAT
GATATGATTTATTACCTTTCAATTATATTTTTCTCAGTTATACTAGAAAATTTCATAATCCTGGGATATATGTAC
CATTGTCAGCTATGACTAAAAATTTGAAAAAGATAAAAATTTCTAGCAAGCCTTTGAAGTTTACCAAGTATAGTC
ACATTCAGTGACAGCCCATTCATTCCAGTAAAGAATCATTTCATTCACTTTGGGAGAGGCCTATAATTACATTTA
TTTGCAATGTTTCTCTTCGCTAGATTGTTACATAGCTCCCATTCTGTTGGTTTTGCTTACAGCATATGGTAACCA
AGGTTAGATGCCAGTTAAAATTCCTTAGAAATTGGATGAGCCTTGAGATTGCTTCTTAACTGGGACATGACATTT
TTCTAGCTCTTATCAAGAATAACAACTTCCACTTTTTTTTAAACTGCACTTTTGACTTTTTTTATGGTATAAAAA
CAATAATTTATAAACATAAAAGCTCATTGTGTTTTTTAGACTTTTGATATTATTTGATACTGTACAAACTTTATT
AAATCAAGATGAAAGACCTACAGGACAGATTCCTTTCAGTGTTCACATCAGTGGCTTTGTATGCAAATATGCTGT
GTTGGACCTGGACGCTATAACTTATTGTAAAGACCTTGGAAATGTGGACATAAGCTCTTTCTTTCCTTTTGTTAC
TGTATTTAGTTTGTGATAAATTTTTCACTGTGTGATATTTATGCTCTAAATCACTACACAAATCCCATATTAAAA
TATACATTGTACCTGAAAAAAAA
```

FIGURE 114

MALERLCSVLKVLLITVLVVEGIAVAQKTQDGQNIGIKHIPATQCGIWVRTSNGGHFASPNYP
DSYPPNKECIYILEAAPRQRIELTFDEHYYIEPSFECRFDHLEVRDGPFGFSPLIDRYCGVKS
PPLIRSTGRFMWIKFSSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIPDCQFELSGADGIVR
SSQVEQEEKTKPGQAVDCIWTIKATPKAKIYLRFLDYQMEHSNECKRNFVAVYDGSSSIENLK
AKFCSTVANDVMLKTGIGVIRMWADEGSRLSRFRMLFTSFVEPPCTSSTFFCHSNMCINNSLV
CNGVQNCAYPWDENHCKEKKKAGVFEQITKTHGTIIGITSGIVLVLLIISILVQVKQPRKKVM
ACKTAFNKTGFQEVFDPPHYELFSLRDKEISADLADLSEELDNYQKMRRSSTASRCIHDHHCG
SQASSVKQSRTNLSSMELPFRNDFAQPQPMKTFNSTFKKSSYTFKQGHECPEQALEDRVMEEI
PCEIYVRGREDSAQASISIDF

Important features of the protein:
Signal peptide:
amino acids 1-22

Transmembrane domain:
amino acids 348-369

N-glycosylation sites.
amino acids 311-315, 385-389, 453-457, 475-479

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 426-430, 479-483

N-myristoylation sites.
amino acids 22-28, 32-38, 54-60, 186-192, 279-285, 318-324, 348-354, 352-358, 441-447

FIGURE 115

GGTCTCTGTCCTTGGCTGTGGCTCCTGCGCTCTGGCTGAGCC**ATG**TTCCTTCTCCTCGCCCTC
CTCACTGAGCTTGGAAGACTGCAAGCCCACGAAGGTTCTGAAGGAATATTTCTGCATGTCACA
GTTCCACGGAAGATTAAGTCAAATGACAGTGAAGTTTCAGAGAGGAAGATGATTTACATCATT
ACAATTGATGGACAACCTTACACTCTACATCTCGGAAAACAATCATTCTTACCCCAGAACTTT
TTGGTTTATACATATAATGAAACTGGATCTTTGCATTCTGTGTCTCCATATTTTATGATGCAT
TGCCATTACCAAGGATATGCTGCCGAATTTCCAAATTCATTTGTGACACTCAGTATATGTTCT
GGTCTCAGGGGATTTCTCCAGTTTGAAAATATCAGTTATGGAATTGAACCAGTAGAATCTTCA
GCAAGATTTGAGCATATAATTTATCAAATGAAAATAATGATCCAAATGTATCCATTTTAGCA
GTAAATTACAGTCATATTTGGCAGAAAGACCAGCCCTACAAAGTTCCTTTAAACTCACAGATA
AAAAATCTTTCAAAACTATTACCCCAATATCTGGAAATATACATTATAGTGGAAAAAGCTTTG
ATGTTTACCCAGTTCAAATTGACTGTTATACTGTCTTCCTTGGAATTGTGGTCAAATGAAAAC
CAGATTTCCACCAGTGGGGATGCTGATGATATATTACAAAGATTTTTGGCATGGAAACGGGAC
TATCTCATCCTACGGCCCCATGACATAGCATACTTACTTGTTTACAGGAAACATCCTAAATAT
GTGGGAGCAACATTTCCTGGCACCGTATGCAATAAAAGCTATGATGCAGGTATTGCTATGTAT
CCAGATGCAATAGGTTTGGAGGGATTTTCGGTTATTATAGCTCAACTGCTTGGCCTTAATGTA
GGATTAACATATGATGACATCACTCAGTGTTTCTGTCTGAGAGCTACATGCATCATGAATCAT
GAAGCAGTGAGTGCCAGTGGTAGAAAGATTTTTAGCAACTGCAGCATGCACGACTATAGATAT
TTTGTTTCAAAATTTGAGACTAAATGCCTTCAGAAGCTTTCAAATTTGCAACCATTACATCAA
AATCAACCAGTGTGTGGTAATGGGATTTTGGAATCCAATGAAGAATGTGACTGTGGTAATAAA
AATGAATGTCAATTTAAGAAGTGCTGTGATTATAACACATGTAAACTGAAGGGCTCAGTAAAA
TGTGGTTCTGGACCATGTTGTACATCAAAGTGTGAGTTGTCAATAGCAGGCACTCCATGTAGA
AAGAGTATTGATCCAGAGTGTGATTTTACAGAGTACTGCAATGGAACCTCTAGTAATTGTGTT
CCTGACACTTATGCACTGAATGGCCGTTTGTGCAAGTTGGGAACTGCCTATTGCTATAACGGA
CAATGTCAAACTACTGATAACCAGTGTGCCAAGATATTTGGAAAAGGTGCTCAAGGTGCTCCA
TTTGCCTGTTTTAAAGAAGTTAATTCTCTGCATGAAAGATCTGAAAACTGTGGTTTTAAAAAT
TCACAACCATTACCTTGTAACGGAAGGATGTTCTCTGTGGAAAATTAGCTTGTGTTCAGCCA
CATAAAAATGCTAATAAAAGTGACGCTCAATCTACAGTTTATTCATATATTCAAGACCATGTA
TGTGTATCTATAGCCACTGGTTCCTCCATGAGATCAGATGGAACAGACAATGCCTATGTGGCT
GATGGCACCATGTGTGGTCCAGAAATGTACTGTGTAAATAAAACCTGCAGAAAAGTTCATTTA
ATGGGATATAACTGTAATGCCACCACAAAATGCAAAGGGAAAGGGATATGTAATAATTTTGGT
AATTGTCAATGCTTCCCTGGACATAGACCTCCAGATTGTAAATTCCAGTTTGGTTCCCCAGGG
GGTAGTATTGATGATGGAAATTTTCAGAAATCTGGTGACTTTTATACTGAAAAAGGCTACAAT
ACACACTGGAACAACTGGTTTATTCTGAGTTTCTGCATTTTTCTGCCGTTTTTCATAGTTTTC
ACCACTGTGATCTTTAAAAGAAATGAAATAAGTAAATCATGTAACAGAGAGAATGCAGAGTAT
AATCGTAATTCATCCGTTGTATCAGAAAGCGATGACGTGGGACAT**TAA**TATTGCACAGAACTT
CCATAGCAAATAACCTAAAGGAACGAATGTGCTTTATTTATAACCTTACGTTATCCCCAATGC
ATTGTAAATGTCAAACTTTTGGAAAATAAAGCCTGCGTGCCCTCCC

FIGURE 116

MFLLLALLTELGRLQAHEGSEGIFLHVTVPRKIKSNDSEVSERKMIYIITIDGQPYTLHLGKQ
SFLPQNFLVYTYNETGSLHSVSPYFMMHCHYQGYAAEFPNSFVTLSICSGLRGFLQFENISYG
IEPVESSARFEHIIYQMKNNDPNVSILAVNYSHIWQKDQPYKVPLNSQIKNLSKLLPQYLEIY
IIVEKALMFTQFKLTVILSSLELWSNENQISTSGDADDILQRFLAWKRDYLILRPHDIAYLLV
YRKHPKYVGATFPGTVCNKSYDAGIAMYPDAIGLEGFSVIIAQLLGLNVGLTYDDITQCFCLR
ATCIMNHEAVSASGRKIFSNCSMHDYRYFVSKFETKCLQKLSNLQPLHQNQPVCGNGILESNE
ECDCGNKNECQFKKCCDYNTCKLKGSVKCGSGPCCTSKCELSIAGTPCRKSIDPECDFTEYCN
GTSSNCVPDTYALNGRLCKLGTAYCYNGQCQTTDNQCAKIFGKGAQGAPFACFKEVNSLHERS
ENCGFKNSQPLPCERKDVLCGKLACVQPHKNANKSDAQSTVYSYIQDHVCVSIATGSSMRSDG
TDNAYVADGTMCGPEMYCVNKTCRKVHLMGYNCNATTKCKGKGICNNFGNCQCFPGHRPPDCK
FQFGSPGGSIDDGNFQKSGDFYTEKGYNTHWNNWFILSFCIFLPFFIVFTTVIFKRNEISKSC
NRENAEYNRNSSVVSESDDVGH

Important features of the protein:
Signal peptide:
amino acids 1-16
Transmembrane domain:
amino acids 665-684
N-glycosylation sites.
amino acids 36-39, 76-79, 122-125, 149-152, 156-159, 177-180, 270-273, 335-338, 441-444, 537-540, 587-590, 601-604, 703-706
Casein kinase II phosphorylation sites.
amino acids 74-77, 208-211, 221-224, 304-307, 337-340, 346-349, 376-380, 415-418, 499-502, 639-642, 708-711
Tyrosine kinase phosphorylation site.
amino acids 243-249
N-myristoylation sites.
amino acids 53-58, 79-84, 266-271, 298-303, 372-377, 403-408, 408-413, 442-447, 462-467, 469-474, 488-493, 567-572, 610-615, 616-621, 634-639
Amidation site.
amino acids 328-331

FIGURE 117

CCCACGCGTCCGCGGACGCGTGGGGCTCAGTGGGCGTCGCGCGAAGGCTAAGGGAGTGTGGCG

GGCGGCTCCGGGAGCCAAC**ATG**CCTCGGTATGCGCAGCTGGTCATGGGCCCCGCGGGCAGCGG

GAAGAGCACCTACTGTGCCACCATGGTCCAGCACTGTGAAGCCCTCAACCGGTCTGTCCAAGT

TGTAAACCTGGATCCAGCAGCAGAACACTTCAACTACTCCGTGATGGCTGACATCCGGGAACT

GATCGAGGTGGATGATGTAATGGAGGATGATTCTCTGCGATTCGGTCCCAACGGAGGATTGGT

ATTTTGCATGGAGTACTTTGCCAATAATTTTGACTGGCTGGAGAACTGTCTTGGCCATGTAGA

GGACGACTATATCCTTTTTGATTGTCCAGGTCAGATTGAGTTGTACACTCACCTGCCTGTGAT

GAAACATCTGGTCCAGCAGCTCGAGCAGTGGGAGTTCCGAGTCTGTGGAGTTTTTCTTGTTGA

TTCTCAGTTCATGGTGGAGTCATTCAAGTTTATTTCTGGCATCTTGGCAGCCCTGAGTGCCAT

GATCTCTCTAGAAATTCCGCAAGTCAACATCATGACAAAAATGGATCTGCTGAGTAAAAAAGC

AAAAAAGGAAATTGAGAAATTTTTAGATCCAGACATGTATTCTTTATTAGAAGATTCTACAAG

TGACTTAAGAAGCAAAAAATTCAAGAAACTGACTAAAGCTATATGTGGACTGATTGATGACTA

CAGCATGGTTCGATTTTTACCTTACGATCAGTCAGATGAAGAAAGCATGAACATTGTATTGCA

GCATATTGATTTTGCCATTCAATATGGAGAAGACCTAGAATTTAAAGAACCAAAGGAACGTGA

AGATGAGTCTTCCTCTATGTTTGACGAATATTTTCAAGAATGCCAGGATGAA**TGA**AGAGTTTA

CTAAAAGTAACCATCTAAAGAGCTTGTGGCCAAACCAGCAGAACATTCTTCTCTTCAAAGGAT

GCAATAGTAGAAAGCTACTTATTTTAATGAAAAAAAGTAAAACTTCGTTCTTTATCAGCCTCA

TGCCTGAATCAAATTTTTAATTATTCTGAAACTGCTGCTGTTTAAAGTGGAATCTTTTAGTAT

TATAACAGCATCACTTTAGATTTTGTAAGTCAAAATTGAAATGAATGCACATAGATTTATATA

TAAATTAGCACCTGAGCTAAGGTTAAGGCCGGTCTAAACTTATTTTCACTTTTTGTATTATTT

TTGAGATGCAGGAATTACTGTAACAAAATATGTATGTCCGAAGGGAAAAAGCTGCAAGGATAT

ATATAAGACCACTGCTTATCTGTATCTTCCCATTTTCCTATATTGAAAATGTATATTATTTAT

ATAACTTAAAAAGTAAAAATAACTATGTTTTGAGAT

FIGURE 118

MPRYAQLVMGPAGSGKSTYCATMVQHCEALNRSVQVVNLDPAAEHFNYSVMADIRELIEVDDV
MEDDSLRFGPNGGLVFCMEYFANNFDWLENCLGHVEDDYILFDCPGQIELYTHLPVMKHLVQQ
LEQWEFRVCGVFLVDSQFMVESFKFISGILAALSAMISLEIPQVNIMTKMDLLSKKAKKEIEK
FLDPDMYSLLEDSTSDLRSKKFKKLTKAICGLIDDYSMVRFLPYDQSDEESMNIVLQHIDFAI
QYGEDLEFKEPKEREDESSSMFDEYFQECQDE

Important features of the protein:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 151-170

N-glycosylation sites.

amino acids 31-35, 47-51

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 212-216

Tyrosine kinase phosphorylation site.

amino acids 189-197

N-myristoylation sites.

amino acids 13-19, 76-82, 154-160

ATP/GTP-binding site motif A (P-loop).

amino acids 10-18

FIGURE 119

```
GGGCGCTGGGAGACACCGGACGCCCGCTCGGCTGCGCTGCGGCTCAGGCCCCCGCTCGGGCCC
GACCCGCTCGGTCACCGCCGGCTCGGGCGCGCACCTGCCGGCTGCGGCCCCAGGGCCATGCGG
AGGCCCACGAGGAGGCCGGCGGCCACGCGCATCCCGTAGCCCAGGTGGCCCAGGTCTGCACCG
CGGCGGCCTCGGCGCCATGGAGCCCCCGTATTCGCTGACGGCGCACTACGATGAGTTCCAAGA
GGTCAAGTACGTGAGCCGCTGCGGCGCGGGGGGCGCGCGCGGGGCCTCCCTGCCCCCGGGCTT
CCCGTTGGGCGCTGCGCGCAGCGTCACCGGGGCCCGGTCCGGGCTGCCGCGCTGGAACCGGCG
CGAGGTGTGCCTGCTGTCGGGGCTGGTGTTCGCCGCCGGCCTCTGCGCCATTCTGGCGGCTAT
GCTGGCCCTCAAGTACCTGGGCCCGGTCGCGGCCGGCGGCGGCGCCTGTCCCGAGGGCTGCCC
TGAGCGCAAGGCCTTCGCGCGCGCCGCTCGCTTCCTGGCCGCCAACCTGGACGCCAGCATCGA
CCCATGCCAGGACTTCTACTCGTTCGCCTGCGGCGGTTGGCTGCGGCGCCACGCCATCCCCGA
CGACAAGCTCACCTATGGCACCATCGCGGCCATCGGCGAGCAAAACGAGGAGCGCCTACGGCG
CCTGCTGGCGCGGCCCGGGGGTGGGCCTGGCGGCGCGGCCCAGCGCAAGGTGCGCGCCTTCTT
CCGCTCGTGCCTCGACATGCGCGAGATCGAGCGACTGGGCCCGCGACCCATGCTAGAGGTCAT
CGAGGACTGCGGGGGGCTGGGACCTGGGCGGCGCGGAGGAGCGTCCGGGGGTCGCGGCGCGATG
GGACCTCAACCGGCTGCTGTACAAGGCGCAGGGCGTGTACAGCGCCGCCGCGCTCTTCTCGCT
CACGGTCAGCCTGGACGACAGGAACTCCTCGCGCTACGTCATCCGCATTGACCAGGATGGGCT
CACCCTGCCAGAGAGGACCCTGTACCTCGCTCAGGATGAGGACAGTGAGAAGATCCTGGCAGC
ATACAGGGTGTTCATGGAGCGAGTGCTCAGCCTCCTGGGTGCAGACGCTGTGGAACAGAAGGC
CCAAGAGATCCTGCAAGTGGAGCAGCAGCTGGCCAACATCACTGTGTCAGAGTATGACGACCT
ACGGCGAGATGTCAGCTCCATGTACAACAAGGTGACGCTGGGGCAGCTGCAGAAGATCACCCC
CCACTTGCGGTGGAAGTGGCTGCTAGACCAGATCTTCCAGGAGGACTTCTCAGAGGAAGAGGA
GGTGGTGCTGCTGGCGACAGACTACATGCAGCAGGTGTCGCAGCTCATCCGCTCCACACCCCA
CCGGGTCCTGCACAACTACCTGGTGTGGCGCGTGGTGGTGGTCCTGAGTGAACACCTGTCCCC
GCCATTCCGTGAGGCACTGCACGAGCTGGCACAGGAGATGGAGGGCAGCGACAAGCCACAGGA
GCTGGCCCGGGTCTGCTTGGGCCAGGCCAATCGCCACTTTGGCATGGCGCTTGGCGCCCTCTT
TGTACATGAGCACTTCTCAGCCGCCAGCAAAGCCAAGGTGCAGCAGCTAGTGGAAGACATCAA
GTACATCCTGGGCCAGCGCCTGGAGGAGCTGGACTGGATGGACGCCGAGACCAGGGCTGCTGC
TCGGGCCAAGCTCCAGTACATGATGGTGATGGTCGGCTACCCGGACTTCCTGCTGAAACCCGA
TGCTGTGGACAAGGAGTATGAGTTTGAGGTCCATGAGAAGACCTACTTCAAGAACATCTTGAA
CAGCATCCCCTTCAGCATCCAGCTCTCAGTTAAGAAGATTCGGCAGGAGGTGGACAAGTCCAC
GTGGCTGCTCCCCCCACAGGCGCTCAATGCCTACTATCTACCCAACAAGAACCAGATGGTGTT
CCCCGCGGGCATCCTGCAGCCCACCCTGTACGACCCTGACTTCCCACAGTCTCTCAACTACGG
GGGCATCGGCACCATCATTGGACATGAGCTGACCCACGGCTACGACGACTGGGGGGGCCAGTA
TGACCGCTCAGGGAACCTGCTGCACTGGTGGACGGAGGCCTCCTACAGCCGCTTCCTGCGAAA
GGCTGAGTGCATCGTCCGTCTCTATGACAACTTCACTGTCTACAACCAGCGGGTGAACGGGAA
ACACACGCTTGGGGAGAACATCGCAGATATGGGCGTCCTCAAGCTGGCCTACCACGCCTATCA
GAAGTGGGTGCGGGAGCACGGCCCAGAGCACCCACTTCCCCGGCTCAAGTACACACATGACCA
GCTCTTCTTCATTGCCTTTGCCCAGAACTGGTGCATCAAGCGGCGGTCGCAGTCCATCTACCT
GCAGGTGCTGACTGACAAGCATGCCCCTGAGCACTACAGGGTGCTGGGCAGTGTGTCCCAGTT
TGAGGAGTTTGGCCGGGCTTTCCACTGTCCCAAGGACTCACCCATGAACCCTGCCCACAAGTG
TTCCGTGTGGTGAGCCTGGCTGCCCGCCTGCACGCCCCACTGCCCCCGCACGAATCACCTCC
TGCTGGCTACCGGGGCAGGCATGCACCCGGTGCCAGCCCCGCTCTGGGCACCACCTGCCTTCC
AGCCCCTCCAGGACCCGGTCCCCCTGCTGCCCCTCACTTCAGGAGGGGCCTGGAGCAGGGTGA
GGCTGGACTTTGGGGGGCTGTGAGGGAAATATACTGGGGTCCCCAGATTCTGCTCTAAGGGGG
CCAGACCCTCTGCCAGGCTGGATTGTACGGGCCCCACCTTCGCTGTGTTCTTGCTGCAAAGTC
TGGTCAATAAATCACTGCACTGTTAAAAAAAAA
```

FIGURE 120

MEPPYSLTAHYDEFQEVKYVSRCGAGGARGASLPPGFPLGAARSVTGARSGLPRWNRREVCLL
SGLVFAAGLCAILAAMLALKYLGPVAAGGGACPEGCPERKAFARAARFLAANLDASIDPCQDF
YSFACGGWLRRHAIPDDKLTYGTIAAIGEQNEERLRRLLARPGGGPGGAAQRKVRAFFRSCLD
MREIERLGPRPMLEVIEDCGGWDLGGAEERPGVAARWDLNRLLYKAQGVYSAAALFSLTVSLD
DRNSSRYVIRIDQDGLTLPERTLYLAQDEDSEKILAAYRVFMERVLSLLGADAVEQKAQEILQ
VEQQLANITVSEYDDLRRDVSSMYNKVTLGQLQKITPHLRWKWLLDQIFQEDFSEEEEVVLLA
TDYMQQVSQLIRSTPHRVLHNYLVWRVVVVLSEHLSPPFREALHELAQEMEGSDKPQELARVC
LGQANRHFGMALGALFVHEHFSAASKAKVQQLVEDIKYILGQRLEELDWMDAETRAAARAKLQ
YMMVMVGYPDFLLKPDAVDKEYEFEVHEKTYFKNILNSIPFSIQLSVKKIRQEVDKSTWLLPP
QALNAYYLPNKNQMVFPAGILQPTLYDPDFPQSLNYGGIGTIIGHELTHGYDDWGGQYDRSGN
LLHWWTEASYSRFLRKAECIVRLYDNFTVYNQRVNGKHTLGENIADMGVLKLAYHAYQKWVRE
HGPEHPLPRLKYTHDQLFFIAFAQNWCIKRRSQSIYLQVLTDKHAPEHYRVLGSVSQFEEFGR
AFHCPKDSPMNPAHKCSVW

Important features of the protein:
Transmembrane domain:
amino acids 64-88
N-glycosylation sites.
amino acids 255-259, 322-326, 656-660
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 722-726
N-myristoylation site.
amino acids 24-30, 26-32, 27-33, 40-46, 47-53, 65-71, 148-154,
169-175, 170-176, 237-243, 450-456, 604-610, 607-613
Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 85-96
Prenyl group binding site.
amino acids 772-777
Neutral zinc metallopeptidases, zinc-binding region signature.
amino acids 609-619
Neutral zinc metallopeptidases, zinc-binding region proteins.
amino acids 609-619

FIGURE 121

CGGACTGCCCGGACCGCGCG**ATG**GAGTCGACCGGCAGCGTCGGGGAGGCCCCGGGCGGACCCC

GGGTGCTGGTGGTGGGCGGCGGCATCGCGGGGCTGGGCGCGGCGCAGAGGCTCTGCGGCCACT

CCGCCTTCCCGCACCTGCGGGTCCTGGAGGCCACGGCCCGCGCCGGGGGCCGCATCCGCTCGG

AGCGCTGCTTCGGTGGCGTGGTGGAGGTGGGCGCGCACTGGATCCATGGGCCCTCCCGGGGTA

ACCCCGTCTTCCAGCTGGCTGCTGAGTACGGGCTGCTGGGGGAGAAGGAGCTGTCCCAGGAGA

ACCAGCTGGTGGAGACCGGGGGTCACGTGGGCCTGCCCTCCGTGAGCTACGCCAGCTCCGGGG

CCAGCGTGAGCCTCCAGCTGGTGGCGGAGATGGCGACTCTGTTCTACGGCCTGATAGACCAGA

CCCGGGAGTTCCTGCACGCTGCAGAGACCCCGGTGCCCAGCGTCGGGGAGTACCTCAAGAAGG

AGATTGGCCAGCACGTGGCCGGCTGGACAGAGGATGAGGAGACCAGGAAGCTGAAGCTGGCCG

TCCTGAACTCCTTCTTCAACCTGGAATGCTGTGTGAGCGGCACCCACAGCATGGACCTGGTGG

CCCTGGCACCCTTTGGGGAGTATACCGTGCTGCCGGGGCTGGACTGCACCTTTTCTAAGGGCT

ATCAAGGACTCACAAACTGCATGATGGCCGCCCTGCCGGAGGACACTGTAGTTTTTGAGAAGC

CTGTGAAGACCATCCACTGGAACGGGTCCTTCCAGGAGGCAGCCTTTCCCGGGGAGACCTTTC

CAGTGTCGGTAGAGTGTGAGGATGGAGACCGGTTCCCGGCGCACCATGTCATCGTCACCGTGC

CCTTAGGTTTTCTTAGGGAACATTTGGACACCTTCTTTGACCCTCCCCTGCCGGCTGAGAAGG

CAGAAGCAATCAGGAAGATAGGCTTTGGGACCAACAACAAAATCTTCCTGGAGTTTGAGGAGC

CCTTCTGGGAGCCAGACTGCCAGCTGATCCAGCTGGTGTGGGAGGACACGTCGCCCCTGGAGG

ATGCTGCCCCTGAGCTACAGGACGCCTGGTTCCGGAAGCTCATTGGCTTTGTGGTCCTGCCTG

CCTTTGCGTCTGTCCACGTTCTCTGTGGGTTCATTGCCGGACTTGAGTCTGAGTTCATGGAGA

CTCTGTCGGATGAAGAAGTACTTCTGTGTCTCACCCAAGTGCTCCGGAGAGTGACAGGAAACC

CACGGCTCCCCGCGCCCAAGAGCGTCCTGCGGTCTCGCTGGCACAGCGCCCCGTACACTAGGG

GGTCCTACAGCTACGTGGCCGTGGGCAGTACTGGGGGCGACCTGGACCTGCTGGCTCAGCCCC

TCCCTGCAGACGGCGCCGGCGCCCAGCTCCAGATCCTGTTTGCGGGGGAAGCCACACATCGCA

CGTTTTACTCCACGACGCACGGGGCTCTGCTGTCGGGATGGAGGGAGGCCGACCGCCTCCTCA

GTCTGTGGGCCCCGCAGGTGCAGCAGCCCAGGCCGAGGCTC**TAG**CTGGGCCCAGCCTACTCTG

TTCCACCCGTGTCGGGGGGTAGGCTGGGACCGTCATTTCTTCTGACAGATTTCAGTCTGGCTTG

AAATTTGGGGATGTTAATGAGGGTCCTCTGGTTTTTGGTAACCAGGGCCACCTTCTCAGTTCT

TGTGTCTGTTATTGGAGTCTGGCCAGGGTTGACTTGAGCTGAGACACCAGATGCTCACGGAGA

TGCTGGACACATAAAGCAAGTTACAGCCACAAAAAAAAAAAA

FIGURE 122

MESTGSVGEAPGGPRVLVVGGGIAGLGAAQRLCGHSAFPHLRVLEATARAGGRIRSERCFGGV
VEVGAHWIHGPSRGNPVFQLAAEYGLLGEKELSQENQLVETGGHVGLPSVSYASSGASVSLQL
VAEMATLFYGLIDQTREFLHAAETPVPSVGEYLKKEIGQHVAGWTEDEETRKLKLAVLNSFFN
LECCVSGTHSMDLVALAPFGEYTVLPGLDCTFSKGYQGLTNCMMAALPEDTVVFEKPVKTIHW
NGSFQEAAFPGETFPVSVECEDGDRFPAHHVIVTVPLGFLREHLDTFFDPPLPAEKAEAIRKI
GFGTNNKIFLEFEEPFWEPDCQLIQLVWEDTSPLEDAAPELQDAWFRKLIGFVVLPAFASVHV
LCGFIAGLESEFMETLSDEEVLLCLTQVLRRVTGNPRLPAPKSVLRSRWHSAPYTRGSYSYVA
VGSTGGDLDLLAQPLPADGAGAQLQILFAGEATHRTFYSTTHGALLSGWREADRLLSLWAPQV
QQPRPRL

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 364-385

N-glycosylation site.

amino acids 253-257

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 408-412

N-myristoylation sites.

amino acids 20-26, 21-27, 25-31, 105-111, 119-125, 164-170, 216-222, 227-233, 443-449, 484-490

Aminooxidase Flavin containing amine oxidase:

amino acids 23-497

FIGURE 123

CGGACGCGTGGGGGAAGATGGATAAATAATTCTGTCACACGTGCCCTGGCCTCTGGAGCTCAGCTGCCAGTCCAC
GTCTAGGGAATCTTAGCATCTGGGACCAAGACACTTTACAGCAATCATCACCCTTTGCAGAGGAGGTGAGCTCAC
CAGGACTCATCTGCCATTTCAGACCTTTTGCTGCTACCTGCCAGGTGGCCCCCACTGCTGACGAGAG**ATG**GTGGA
TCTCTCAGTCTCCCCGGACTCCTTGAAGCCAGTATCGCTGACCAGCAGTCTTGTCTTCCTCATGCACCTCCTCCT
CCTTCAGCCTGGGGAGCCGAGCTCAGAGGTCAAGGTGCTAGGCCCTGAGTATCCCATCCTGGCCCTCGTCGGGGA
GGAGGTGGAGTTCCCGTGCCACCTATGGCCACAGCTGGATGCCCAGCAAATGGAGATCCGCTGGTTCCGGAGTCA
GACCTTCAATGTGGTACACCTGTACCAGGAGCAGCAGGAGCTCCCTGGCAGGCAGATGCCGGCGTTCCGGAACAG
GACCAAGTTGGTCAAGGACGACATCGCCTATGGCAGCGTGGTCCTGCAGCTTCACAGCATCATCCCCTCTGACAA
GGGCACATATGGCTGCCGCTTCCACTCCGACAACTTCTCTGGCGAAGCTCTCTGGGAACTGGAGGTAGCAGGGCT
GGGCTCAGACCCTCACCTCTCCCTTGAGGGCTTCAAGGAAGGAGGCATTCAGCTGAGGCTCAGATCCAGTGGCTG
GTACCCCAAGCCTAAGGTTCAGTGGAGAGACCACCAGGGACAGTGCCTGCCTCCAGAGTTTGAAGCCATCGTCTG
GGATGCCCAGGACCTGTTCAGTCTGGAAACATCTGTGGTTGTCCGAGCGGGAGCCCTCAGCAATGTGTCCGTCTC
CATCCAGAATCTCCTCTTGAGCCAGAAGAAAGAGTTGGTGGTCCAGATAGCAGACGTGTTCGTACCCGGAGCCTC
TGCGTGGAAGAGCGCGTTCGTCGCGACCCTGCCGCTGCTGTTGGTCCTCGCGGCGCTGGCGCTGGGCGTCCTCCG
GAAGCAGCGGAGAAGCCGAGAAAAGCTGAGGAAGCAGGCGGAGAAGAGACAAGAGAAACTCACTGCAGAGCTGGA
AAAGCTTCAGACAGAGCTTGACTGGAGACGGGCTGAAGGCCAGGCTGAGTGGAGAGCAGCCCAAAAATATGCAGT
GGATGTGACGCTGGACCCGGCCTCGGCGCACCCCAGCCTGGAGGTGTCGGAGGATGGCAAGAGCGTGTCTTCCCG
CGGGGCGCCGCCAGGCCCGGCGCCTGGCCACCCGCAGCGGTTCTCGGAGCAGACGTGCGCGCTGAGCCTGGAGCG
GTTCTCCGCGGCCGCCACTACTGGGAGGTGCACGTGGGCCGCCGCAGCCGCTGGTTCCTGGGCGCCTGCCTGGC
CGCGGTGCCGCGCGCGGGGCCTGCGCGCGTGAGCCCTGCGGCCGGCTACTGGGTGCTGGGGCTGTGGAACGGCTG
CGAGTACTTCGTCCTGGCCCCGCACCGCGTCGCGCTCACCCTGCGCGTGCCCCCGCGCGGCGCCTGGGCGTCTTCCT
GGACTACGAGGCCGGAGAGCTGTCCTTCTTCAACGTGTCCGACGGCTCCCACATCTTCACCTTCCACGACACCTT
CTCGGGCGCGCTCTGTGCGTACTTCAGGCCCAGGGCCCACGACGGCGGCGAACATCCGGATCCCCTGACCATCTG
CCCGCTGCCGGTTAGAGGGACGGGCGTCCCCGAAGAGAACGACAGTGACACCTGGCTACAGCCCTATGAGCCCGC
GGACCCCGCCCTGGACTGGTGG**TGA**GGCGCCCTCGTGGCCGCGGGACTGGCCCCGGGGGGGCCCCCTGGATCCCAG
GCCAGCGCTTTGCTCTCCTGCTCCGTCTGAAGGGAGCAGGTGCACCAGCCAAAATGTCAGCGAGGGGGACAAAGA
GAGGGACCTTTGCCTACGTAGATGTGTATGTGTAGTGCGATTTTCTTCAAGGAAAGGAGACAAGTCCAAAGCTCG
TTTGTGGATTGTGGGACTGAGCGAAGGAGTACAAATATATCCACGTCGCTCAGAGCTGGGGTGCTCACGGTGGGC
GGTGGGCAAGAAGCCAGCATGGAAGAAAGAAGGGAGAAAACTTTGGTGACTGCCTTAGAGGGATCAGTTAATTTG
TATAGTTTTATATTTTTTGTATATGTTTGCTAGCTCTAAAAAGGTCGAGATGCAATAACACTTCGTAAGCAACGA
GTTCACCTAAGTAAGGCTCAGATCCTAGTTTTAAAAACCATTTCCCATTAAAATGAAGTTGGAGGAACAGCTGCT
TCTGAGCCGGGGCAAAAATTTCAAGGTGAGCCTGGAGCATTGTGTGTGGTGAAGTAAAATAAAGGCTCAAAACGT
GACGGCAACCCGGCAAAAGGGTAGGGAGCCAGGCCGAAGGGGCCTCACTGACCAATTGTGGGACAATTTGAACAT
CAGGATGAATAATGACAGGAGAGATTATAACACACTGAATAAAAACATAATCCATGAGTTCATGCTGATACTCAA
ATTTCTTTTTAAAAAGGAGAAACAGGAAGGTTTCTTTTGGAGGTGAAATCTAATTATTGGTGAGAGTCTTGGAGA
ACAGGCTGTTTCCAGTCTCAAAGCAGTAACCTTATACACTACTTATAAGTTTGAAAGGGGAAAGGTTACCTTTAC
AATGGAGACATCTACCAGATCATCCAAGTGATTAAATTTAACATCATCAATGATGGGACCAAGGACATTATTAGT
TTGACAACTGGGGAAAGAAGTGTTCTTCACCCCCCTACCCCCAAGACATTCTCTCTGTCGGCCAGGCTGGAGTGCA
GCCTCAACCTCCTGGGCCCAAGTGATCCTCCCACCTCAGCACACAACACCATGCCCAATTTTAAGTGCGTTATAG
AGACGGGGGTCTCACTTTGTTACCCAGGCTGGTCTCAAACTCCTGCGCTCAAGCAATCCTCCCACCTGGGCCTCC
CAAAATGCTGGGTGTACAGGCATGAGCCGCTGTGCCTGGCTTCATTTTCAGAGTGAGACATTTGTACTGTGGCTA
TGTAGGAGAACATTCTTGTTCTTAGCAAACATACTGAAGTTTTTAGATATTAATTACCACAGTGTCTGCCACTGA
ATTTCCAGTGACTAAGTGGAAAAATATAAAACATATGAATATAAAGAAAGAAAGAGACAAGTCAAATGTAGTAAA
ATGACAACACTTGGTGACTCTAGGTGACTGGTCGACAGATGTTCATTGTACTATCAATGTGGCTTTGCTGTGGGT
TTGAAATTTTGCAAACTAAGAGTTGGGTGGCGGGGAGAAGGATACACCAAAAAACTAAGTGATTATCTTTGGATG
GGAAAATGTTTGGTAATTGCATTCTTAAAATGTCTTCTTTGTATTTTTTAATGTTCAATAATGTATATGTATCAG
TTCTGTAATAAAGGGGAAAACACTTTTCA

FIGURE 124

MVDLSVSPDSLKPVSLTSSLVFLMHLLLLQPGEPSSEVKVLGPEYPILALVGEEVEFPCHLWP
QLDAQQMEIRWFRSQTFNVVHLYQEQQELPGRQMPAFRNRTKLVKDDIAYGSVVLQLHSIIPS
DKGTYGCRFHSDNFSGEALWELEVAGLGSDPHLSLEGFKEGGIQLRLRSSGWYPKPKVQWRDH
QGQCLPPEFEAIVWDAQDLFSLETSVVVRAGALSNVSVSIQNLLLSQKKELVVQIADVFVPGA
SAWKSAFVATLPLLLVLAALALGVLRKQRRSREKLRKQAEKRQEKLTAELEKLQTELDWRRAE
GQAEWRAAQKYAVDVTLDPASAHPSLEVSEDGKSVSSRGAPPGPAPGHPQRFSEQTCALSLER
FSAGRHYWEVHVGRRSRWFLGACLAAVPRAGPARLSPAAGYWVLGLWNGCEYFVLAPHRVALT
LRVPPRRLGVFLDYEAGELSFFNVSDGSHIFTFHDTFSGALCAYFRPRAHDGGEHPDPLTICP
LPVRGTGVPEENDSDTWLQPYEPADPALDWW

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 247-272

N-glycosylation sites.

amino acids 102-106, 139-143, 224-228, 464-468, 516-520

Tyrosine kinase phosphorylation site.

amino acids 105-114

N-myristoylation sites.

amino acids 129-135, 220-226, 399-405, 423-429, 480-486

Amidation site.

amino acids 390-394

FIGURE 125

```
TATAGTCCCAGCTACTCATGGGGCTGATGCAGGTTGAGGCAGGAGGTTCATGAGCCCAGGAGGTTGGAGCTGTAA
TGAGCTAGGATTCTGCCTCTGCACTCCTAGCTGGATGACAGAGCAAGACCCTGTCTCAAAAAAGAAAAAAAAAAA
AAAAAGAATGCATGAACCAGACATGACAGTTCCTGGCCTCAAAGATCTTCCAAAGGAAATGATTTTTTTTTAACC
ACCAATGCTGCAGGAAAAAGCAACATATTTAAGTTATCCAATAACACCTATCCAATAATTGTAAATCATTATCAT
GAC**ATG**GTAGAGTTGTTTATATTTCTTTTCCTTTTAGGTGAAACACCATTCAAAGTCGTAGTCAAATCTCTTTCA
CCTAAAGAGTTGGTCCGGATACATGTCCCTAAACCTTTGGACAGGAATGATGGAACATTTTTGATGAGATATAGG
ATGTATGAAACTGTCGATGAAGGCCTGAAGATAGAGGTCCTTTATGGTGATGAACATGTGGCTCAGTCTCCCTAT
ATTTTGAAAGGACCAGTGTACCATGAGTACTGTGAGTGTCCGGAAGATCCTCAGGCCTGGCAGAAGACTCTTTCT
TGTCCAACCAAGGAACCACAGATTGCAAAGATTTTGCTTCCTTTCCCAGCATCAATCTCCAGCAAATGCTAAAA
GAAGTCCCCAAAAGGTTTGGGGATGAGAGAGGTGCCATTGTTCATTACACGATTCTCAATAACCATGTTTACCGG
AGATCTTTAGGGAAATACACAGACTTCAAGATGTTCTCTGATGAGATTTTGTTATCATTGACAAGAAAGGTCCTT
CTCCCAGATTTAGAATTTTATGTTAATCTTGGAGATTGGCCCTTGGAGCATCGAAAAGTCAATGGAACCCCTAGC
CCCATACCTATCATTTCATGGTGTGGCTCTCTGGATTCAAGAGATGTTGTCCTTCCAACGTATGACATCACCCAC
TCCATGCTTGAAGCCATGCGGGGTGTTACAAATGATCTCCTCTCTATTCAGGGAAATACAGGGCCTTCCTGGATC
AATAAAACAGAGAGAGCTTTCTTCAGAGGTAGAGACAGCCGAGAGGAGAGGCTCCAGTTGGTACAGCTGTCCAAA
GAAAATCCTCAGCTACTAGATGCAGGAATTACAGGATATTTCTTTTTCCAAGAGAAAGAAAAGGAGCTTGGAAAA
GCCAAGTTGATGGGTTTCTTTGATTTCTTTAAGTACAAGTATCAAGTAAATGTGGATGGGACCGTGGCTGCTTAC
AGATATCCATATCTCATGCTGGGCGACAGTCTGGTTTTAAAGCAGGACTCGCCATATTATGAACATTTCTACATG
GCACTAGAACCTTGGAAGCATTATGTTCCAATTAAAAGAAATCTGAGTGATTTATTAGAGAAAGTTAAATGGGCT
AAGGAAAATGATGAAGAAGCCAAGAAGATTGCAAAGAAGGACAGTTGATGGCTAGGGACCTACTACAGCCACAC
AGGCTTTACTGCTACTATTACCAAGTACTGCAGAAATATGCCGAGCGCCAGTCCAGCAAACCCGAAGTACGTGAT
GGAATGGAACTTGTTCCTCAGCCAGAAGATAGCACAGCCATCTGCCAGTGCCACAGGAAAAAGCCTTCAAGAGAA
GAACTT**TGA**GTCAGCCCAGAATCACACTCCTGTGTATCCCGGCTACACTTTAAGGAAAGATTGAATCTAAGCTGT
GAAGGACAGTATAGAAGACTGCACCAAGTGGACTAGTTCTCCCGGTGGCTTTATATATGTAGATGGATATAGCAG
TACTGGTTGAGTATCCCTCATCTGAAATGCTTAGGACCAGGAGTGTTTCAGGCTTCAGATTTTTTAAGATTTGGG
AATATTTGCATGTACATAATGAGGTATCTTGGGGATGAGATCCAAGTCTAAACACAAAATTCATTTATATTTTAT
ATATACCTTGTTCACATACCCTGAAGGTAATTTTATATAATATTTTTAATAATTTGTGCATGAAACAAAGTTTGT
ATACATTGAACTGTCAGAAAGCAAAGGTGTCACTATCTTAGCGACCCAAGTGGTGGTGTCAGCACTCAAAAAGTT
TTGGATTTTGGGGTATTTCAGATTTTAGATTTTTGTATGAGGAATGTTCAACCTGTATTTGAACAAGCATTACCA
AATATCATTGAATATTAATATCTTTTGCGTAAAAACTGCTATTATCAGCATCATAGTTTCTCTAAAAAGAAAACT
TGGGGATCATAGCCGATAGAGAGACTTGCTAAAATATAAATCAGCCTCTGCAAAACTGTTTACATATTTATTGGT
TTACATATTTTATTGGTTTATTTCTATCCCCTGTTCACTTTTTCTCTTCCACTTCCAATTATGAAGAGAAAATAT
TTGTTCAGGGTTGTCCCCCCGCCCCCCGTCACTGCATAATTTCTCCTCTTACAAGCTGCTTTTGGCTTTCATTAA
TAACAGCTTCCTTTTAGAAGGTCTGATAAGGATATTTAAGGAAGAAGAGAATGACTCTGTTATTAAAGGTGGCAT
GGAGACTGTGGAGGGAATATTTTTTAAAGCACTACTCATATCCTTTAAACTAAATTTTGCCAAAGCCCGAGACAA
CATTAAGGAGAAATTGTACCTTAAGTTAGTAATTCCAAATCTATCTGAGTTGTATACCCATCAAAGACAATACAG
TTATTAACATAGATGAAGGTATGCTATAGGCATCATTCATTATCTCTATATTGAATAGGTGAAAGATAACTGTAG
TCAGGTGAAAGGCATTCATCATTTTTAAGCTGAAAAGGGGATCCTTGAAAACACTGAAAACCTCTACAACAATCT
TCAGGAAGCCTGCTATCTTGGGATTCACTAATAATAGGCCAAGAACAAAGGCAAGCATCCATTCCTCACTCCACC
ACTTTTCTATTTCAGTGGGTGTCATTGCTACGATGAAGACTTTGGAAATTTCCTTTCTCTTTTAGGACAGGGTCA
GGATTTAGGACTCATAGCCTGAAAGCTCATTACATACTCCTTGTAACCATCAGTCCAAGGTTCAGTTCACTAAAG
TGCATGTTCTAAAACAAGAGCTATCCTCATTCCAAATTTTAAAATATGTACTCTGGCCGGTTGCAGTGGCTCACG
CCTGTAATCCCAGCACTTTGGCAGGCCGAGATGGGCGGATCTTTTGAGGTCAGGAGTTTGAGACCAGCCTGGCCA
ACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGGCATGGTGGCATTTGCCTGTAATCCCAGCT
ACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTACACCACTG
CACTCCAGCCTGGGTGACAGAGTGAGACTCCATCTCAAAAACTGAAAATAAAAATAAAAATATGTATTCTCCTAA
CTGAAATATTTACTTAATCTGGAAAACAATGTAACTATTTTTAAAGTGGTTACATCTATTCTTGCTGAAGAACAA
TAAACAGAATTTTTTGACTAAGCATAACCAAATTTCAGAACAGTCTAATCAATGCCAAGTATCCAAGGCAAACTC
TAATACCCATCCATTGTGCAAAACCACAAGCACGCAAGTATTAAATAAGAGCAAGCTGTCCTGAGCCCATACCTA
ATGAATTTGTGTCTTAAATATTGTACATTGTGTTTGAGGCTTGTCAAAACTGGGATTATGGCAAGAAAGGTTGCC
TAACTCATACCTTTCTGCCTCAAATTCCAGGTGCTAAAGGCTAATGGCATTTTAAACATCTTACATTTTTAAAAA
TTTATATTGCCTCTGCCAAACAGGCCTAATAGTTAAAAGCAAGTTGAGACAAACCAGGCAGATTCAGTGTGTGGA
ACAGGAAGGATGTGCTTTAAAAAAAGGTGGAATCCCTCAAAAAATTCTATAGGGAGACAGCAGCCTTAATCTACA
TAATTCTTCATCTCGCCAATTCAGCCGCAGCCTTTAAAGAGTTAGTGTTAATGGCTTTCTGGTTTGAAAACAAAA
ATGCATCTATGTGGTTGAAAGTTTGGGAGGAGATTCACCAATATCTGAGGAGAAGATGGAGTGAAGGGAATTCTT
ACTTTTTGCTTTATACCTTTCTATAATATTTAGATTTTTTTTTACTGTAAGTATGGATCAAATTGCAAAATAAAG
AAAAATGCCAACCTTAGAAAAGACAATAAATGCACAAAAGATATAAACAGGAACAGCAAATATTTATATTTTTTC
CATTTTGCTCTTTTTAAATCTATGTTTAGAACTTTATATCTTGGGACTTATGTATATATATACCTTTTAAATAAA
ATAAATTTTCTAAATAAAAAGTTG
```

FIGURE 126

MVELFIFLFLLGETPFKVVVKSLSPKELVRIHVPKPLDRNDGTFLMRYRMYETVDEGLKIEVL
YGDEHVAQSPYILKGPVYHEYCECPEDPQAWQKTLSCPTKEPQIAKDFASFPSINLQQMLKEV
PKRFGDERGAIVHYTILNNHVYRRSLGKYTDFKMFSDEILLSLTRKVLLPDLEFYVNLGDWPL
EHRKVNGTPSPIPIISWCGSLDSRDVVLPTYDITHSMLEAMRGVTNDLLSIQGNTGPSWINKT
ERAFFRGRDSREERLQLVQLSKENPQLLDAGITGYFFFQEKEKELGKAKLMGFFDFFKYKYQV
NVDGTVAAYRYPYLMLGDSLVLKQDSPYYEHFYMALEPWKHYVPIKRNLSDLLEKVKWAKEND
EEAKKIAKEGQLMARDLLQPHRLYCYYYQVLQKYAERQSSKPEVRDGMELVPQPEDSTAICQC
HRKKPSREEL

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 250-254, 363-367

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 444-448

N-myristoylation site.

amino acids 208-214, 319-325, 388-394

Endoplasmic reticulum targeting sequence.

amino acids 448-453

Mitochondrial energy transfer proteins signature.

amino acids 25-34

FIGURE 127

AGCCGTCGGAGGGAGCCGGAGCGCTTCTCCCGAGTTGGTGATAGATTGGTGGTCATCCAACAT
GCAGAAATGAATGAGCAGTGAAAAGCAGCAGAGCCGATGGGTCATGAGGATGTAAGTGCGTTT
GAAGGCTTCCACACCCTCTACTCCAGGAATCATGAATAAACTGGAGGATAAGCAGGACCAGAT
GATACC**ATG**AAGAGAAGTTTACAGGCCCTCTATTGCCAACTGTTAAGTTTCCTGCTGATCTTG
GCACTGACCGAAGCGCTGGCATTTGCCATCCAGGAACCATCTCCCAGGGAATCTCTTCAGGTC
CTCCCTTCAGGCACTCCCCCGGGAACCATGGTGACAGCACCCCACAGCTCTACCAGACATACT
TCTGTGGTGATGCTGACCCCCAATCCCGATGGACCCCCCTCACAGGCTGCAGCTCCCATGGCA
ACACTGACACCCCGTGCAGAGGGGCACCCTCCTACGCACACCATCTCCACCATCGCTGCGACA
GTAACCGCCCCCTATTCTGAAAGCTCCCTGTCCACAGGGCCCGCTCCAGCAGCCATGGCAACC
ACATCCTCCAAGCCAGAGGGCCGCCCTCGAGGGCAGGCTGCCCCCACCATCCTGCTGACAAAG
CCACCGGGGGCCACCAGCCGCCCCACCACAGCGCCCCCCGCACTACCACACGCAGGCCCCCC
AGGCCCCCAGGCTCTTCCCGAAAAGGGGCTGGTAATTCATCACGCCCTGTCCCGCCTGCACCT
GGTGGCCACTCCAGGAGTAAAGAAGGACAGCGAGGACGAAATCCAAGCTCCACACCTCTGGGG
CAGAAGCGGCCCCTGGGGAAAATCTTTCAGATCTACAAGGGCAACTTCACAGGGTCTGTGGAA
CCAGAGCCCTCTACCCTCACCCCCAGGACCCCACTCTGGGGCTACTCCTCTTCACCACAGCCC
CAGACAGTGGCTGCGACCACAGTGCCCAGCAATACCTCATGGGCACCACCACCACCTCCCTG
GGGCCTGCAAAGGACAAGCCAGGCCTTCGCAGAGCAGCCCAGGGGGGTGGTTCTACCTTCACC
AGCCAAGGAGGGACACCAGATGCCACAGCAGCCTCAGGTGCCCCTGTCAGTCCACAAGCTGCC
CCAGTGCCTTCTCAGCGCCCCCACCACGGTGACCCACAGGATGGCCCCAGCCATAGTGACTCT
TGGCTTACTGTTACCCCTGGCACCAGCAGACCTCTGTCTACCAGCTCTGGGGTCTTCACGGCT
GCCACGGGGCCCACCCCAGCTGCCTTCGATACCAGTGTCTCAGCCCCTTCCCAGGGGATTCCT
CAGGGAGCATCCACAACCCCACAAGCTCCAACCCATCCCTCCAGGGTCTCAGAAAGCACTATT
TCTGGAGCCAAGGAGGAGACTGTGGCCACCCTCACCATGACCGACCGGGTGCCCAGTCCTCTC
TCCACAGTGGTATCCACAGCCACAGGCAATTTCCTCAACCGCCTGGTCCCCGCCGGGACCTGG
AAGCCTGGGACAGCAGGGAACATCTCCCATGTGGCCGAGGGGGACAAACCGCAGCACAGAGCC
ACCATCTGCCTGAGCAAGATGGATATCGCCTGGGTGATCCTGGCCATCAGCGTGCCCATCTCC
TCCTGCTCTGTCCTGCTGACGGTGTGCTGCATGAAGAGGAAGAAGAAGACCGCCAACCCGGAG
AACAACCTGAGCTACTGGAACAACACCATCACCATGGACTACTTCAACAGGCATGCTGTGGAG
CTGCCCAGGGAGATCCAGTCCCTTGAAACCTCTGAGGACCAGCTCTCAGAGCCCCGCTCCCCA
GCCAATGGCGACTATAGAGACACTGGGATGGTCCTTGTTAACCCCTTCTGTCAAGAAACACTG
TTTGTGGGAAACGATCAAGTATCTGAGATC**TAA**CTACAGCAGGCATCACTTTGCCATTCCGTA
TTTTTCGTCTCTAAATTATAAATATACAAATATATATATTATAAATATAACCTTGTGTAACCC
TGACTTAATGAGAAACATTTTCAGCTTTTTTTCCTATGAATTGTCAACATCTTTTTTACAAGT
GTGGTTTAAAAAAAAAAAAAACTTTACAGAATGATCTGTGGCTTTATAAAATAAAGGTATTTCT
AAGCAAAAAAAAAAAAAAAAA

FIGURE 128

MKRSLQALYCQLLSFLLILALTEALAFAIQEPSPRESLQVLPSGTPPGTMVTAPHSSTRHTSV
VMLTPNPDGPPSQAAAPMATLTPRAEGHPPTHTISTIAATVTAPYSESSLSTGPAPAAMATTS
SKPEGRPRGQAAPTILLTKPPGATSRPTTAPPRTTTRRPPRPPGSSRKGAGNSSRPVPPAPGG
HSRSKEGQRGRNPSSTPLGQKRPLGKIFQIYKGNFTGSVEPEPSTLTPRTPLWGYSSSPQPQT
VAATTVPSNTSWAPTTTSLGPAKDKPGLRRAAQGGGSTFTSQGGTPDATAASGAPVSPQAAPV
PSQRPHHGDPQDGPSHSDSWLTVTPGTSRPLSTSSGVFTAATGPTPAAFDTSVSAPSQGIPQG
ASTTPQAPTHPSRVSESTISGAKEETVATLTMTDRVPSPLSTVVSTATGNFLNRLVPAGTWKP
GTAGNISHVAEGDKPQHRATICLSKMDIAWVILAISVPISSCSVLLTVCCMKRKKKTANPENN
LSYWNNTITMDYFNRHAVELPREIQSLETSEDQLSEPRSPANGDYRDTGMVLVNPFCQETLFV
GNDQVSEI

Important features of the protein:
Signal peptide:
amino acids 1-28

Transmembrane domain:
amino acids 469-487

N-glycosylation sites.
amino acids 178-182, 223-227, 261-265, 446-450, 504-508, 509-513

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 495-499

N-myristoylation sites.
amino acids 44-50, 48-54, 175-181, 222-228, 279-285, 286-292,
288-294, 296-302, 351-357, 374-380, 427-433, 442-448

TonB-dependent receptor proteins signature 1.
amino acids 1-44

FIGURE 129

AGGCGAGGCGCGGCGCCGCTGCACACACGCACACGGAGCT**ATG**GGGTGCCATGTTGCCACCAG
CTGCCACGTGGCCTGGCTTTTGGTGCTGATCTCTGGATGCTGGGGCCAGGTGAACCGGCTGCC
CTTCTTCACCAACCACTTCTTTGATACATACCTGCTGATCAGCGAGGACACGCCTGTGGGTTC
TTCTGTGACCCAGTTGCTGGCCCAAGACATGGACAATGACCCCCTGGTGTTTGGCGTGTCTGG
GGAGGAGGCCTCTCGCTTCTTTGCAGTGGAGCCTGACACTGGCGTGGTGTGGCTCCGGCAGCC
ACTGGACAGAGAGACCAAGTCAGAGTTCACCGTGGAGTTCTCTGTCAGCGACCACCAGGGGGT
GATCACACGGAAGGTGAACATCCAGGTCGGGGATGTGAATGACAACGCGCCCACATTTCACAA
TCAGCCCTACAGCGTCCGCATCCCTGAGAATACACCAGTGGGGACGCCCATCTTCATCGTGAA
TGCCACAGACCCCGACTTGGGGGCAGGGGGCAGCGTCCTCTACTCCTTCCAGCCCCCCTCCCA
ATTCTTCGCCATTGACAGCGCCCGCGGTATCGTCACAGTGATCCGGGAGCTGGACTACGAGAC
CACACAGGCCTACCAGCTCACGGTCAACGCCACAGATCAAGACAAGACCAGGCCTCTGTCCAC
CCTGGCCAACTTGGCCATCATCATCACAGATGTCCAGGACATGGACCCCATCTTCATCAACCT
GCCTTACAGCACCAACATCTACGAGCATTCTCCTCCGGGCACGACGGTGCGCATCATCACCGC
CATAGACCAGGATAAAGGACGTCCCCGGGGCATTGGCTACACCATCGTTTCAGGGAATACCAA
CAGCATCTTTGCCCTGGACTACATCAGCGGAGTGCTGACCTTGAATGGCCTGCTGGACCGGGA
GAACCCCCTGTACAGCCATGGCTTCATCCTGACTGTGAAGGGCACGGAGCTGAACGATGACCG
CACCCCATCTGACGCTACAGTCACCACGACCTTCAATATCCTGGTTATTGACATCAATGACAA
TGCCCCGGAGTTCAACAGCTCCGAGTACAGCGTGGCCATCACTGAGCTGGCACAGGTCGGCTT
TGCCCTTCCACTCTTCATCCAGGTGGTGGACAAGGATGAGAATTTGGGCCTGAACAGCATGTT
TGAGGTGTACTTGGTGGGGAACAACTCCCACCACTTCATCATCTCCCCGACCTCCGTCCAGGG
GAAGGCGGACATTCGTATTCGGGTGGCCATCCCACTGGACTACGAGACCGTGGACCGCTACGA
CTTTGATCTCTTTGCCAATGAGAGTGTGCCTGACCATGTGGGCTATGCCAAGGTGAAGATCAC
TCTCATCAATGAAAATGACAACCGGCCCATCTTCAGCCAGCCACTGTACAACATCAGCCTGTA
CGAGAACGTCACCGTGGGGACCTCTGTGCTGACAGTCCTGGTGAGTCCCCGCTTCACTGCAGG
GCCACTGAGCTCTCCAGGGCCGACTGTGGTGAGGCACCCAGAGGGATTTTGTCCAAGGGACCT
CAGCAATCAGGGAAGGAGGCACCCCCAAATCCCTGAGCTGTGTTTGTTGGTGTAT**TAA**ATAAA
GTTTTTGGACTCTTCAGGAAGGGGCTCCCTTGACCTAGGTTGCAATATGGAAAAGGAGCCAAC
CTGAGGGGTGACGAGACTGAGCTGAGGACACTGGTTTTCTGCCTTTCCCTGAGAGAGACTCAG
TGAGGGTGGGCTGGGAGCCCTGGAAGCCCCCTCAAATGGGTGGGAAGGTGCCAGCCATCCTTG
AGAAGGGCAACCCTCTCCATGTGAGCACAGGCACCAGAGAGGGGCAGGCGCCTGGAGGGTACC
GGGGCACCCCCAGCTGCCCATGGCTGGACTTGCCCTTTGACAAGGGGCCCTCCCAGTGTCATT
TGTATCTGTCAGTACTCTTGGTTGCAAGGGACAGAAACCCTTAAGTAGTTCAAGCAAAAAAGG
ATTGGCTCATGTAACTCAAAAGTATAAGTGATTTCAGGCCGGGCTCGGTGGCTCACGCCTGTC
ATCCAACACCTTGAGAAAGCCGAGGTGGGCGGATCACTTGAGGTCGGGAGTTTGAGACCAGCC
TGGCCAACATGGCAAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGGTGTGGTGGCAC
ACGCCTGTAGTCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGCGG
AGGTTGCAGTGAGCCGAGATTGTGTCACTGCCCTCCAGCCTGGGCGACAGAGCCAGATTCTGT
CTC

FIGURE 130

MGCHVATSCHVAWLLVLISGCWGQVNRLPFFTNHFFDTYLLISEDTPVGSSVTQLLAQDMDND
PLVFGVSGEEASRFFAVEPDTGVVWLRQPLDRETKSEFTVEFSVSDHQGVITRKVNIQVGDVN
DNAPTFHNQPYSVRIPENTPVGTPIFIVNATDPDLGAGGSVLYSFQPPSQFFAIDSARGIVTV
IRELDYETTQAYQLTVNATDQDKTRPLSTLANLAIIITDVQDMDPIFINLPYSTNIYEHSPPG
TTVRIITAIDQDKGRPRGIGYTIVSGNTNSIFALDYISGVLTLNGLLDRENPLYSHGFILTVK
GTELNDDRTPSDATVTTTFNILVIDINDNAPEFNSSEYSVAITELAQVGFALPLFIQVVDKDE
NLGLNSMFEVYLVGNNSHHFIISPTSVQGKADIRIRVAIPLDYETVDRYDFDLFANESVPDHV
GYAKVKITLINENDNRPIFSQPLYNISLYENVTVGTSVLTVLVSPRFTAGPLSSPGPTVVRHP
EGFCPRDLSNQGRRHPQIPELCLLVY

Important features of the protein:
Signal peptide:
amino acids 1-23
Transmembrane domain:
amino acids 355-374
N-glycosylation sites.
amino acids 155-159, 206-210, 349-353, 393-397, 434-438, 466-470,
472-476
N-myristoylation sites.
amino acids 2-8, 49-55, 162-168, 270-276, 278-284, 316-322
Amidation site.
amino acids 515-519
Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 11-22
Leucine zipper pattern.
amino acids 298-320
PTS HPR component serine phosphorylation site signature.
amino acids 377-393
Cadherins extracellular repeated domain signature.
amino acids 120-131, 336-347
Cadherins extracellular
amino acids 120-144, 336-360

FIGURE 131

GTGGGCCGCCCCTGCTGCTGCCGTCC**ATG**CTGATGTTTGCGGTGATCGTGGCCTCCAGCGGGC

TGCTGCTCATGATCGAGCGGGGCATCCTGGCCGAGATGAAGCCCCTGCCCCTGCACCCGCCCG

GCCGCGAGGGCACAGCCTGGCGCGGGAAAGCCCCCAAGCCTGGGGGCCTGTCCCTCAGGGCTG

GGGACGCGGACTTGCAAGTGCGGCAGGACGTCCGGAACAGGACCCTGCGGGCGGTGTGCGGAC

AGCCAGGCATGCCCCGGGACCCCTGGGACTTGCCGGTGGGGCAGCGGCGCACCCTGCTGCGCC

ACATCCTCGTAAGTGACCGTTACCGCTTCCTCTACTGCTACGTCCCCAAGGTGGCCTGCTCTA

ACTGGAAGCGGGTGATGAAGGTGCTGGCAGGCGTCCTGGACAGCGTGGACGTCCGCCTCAAGA

TGGACCACCGCAGTGACCTGGTGTTCCTGGCCGACCTGCGGCCTGAGGAGATTCGCTACCGCC

TGCAGCACTACTTTAAGTTCCTGTTTGTGCGGGAGCCCTTGGAACGCCTCCTCTCTGCCTACC

GCAACAAGTTTGGCGAGATCCGAGAGTACCAGCAACGCTATGGGGCTGAGATAGTGAGGCGGT

ACAGGGCTGGAGCGGGGCCCAGCCCTGCAGGCGACGATGTCACATTCCCCGAGTTCCTGAGAT

ACCTGGTGGATGAGGACCCTGAGCGCATGAATGAGCATTGGATGCCCGTGTACCACCTGTGCC

AGCCTTGTGCCGTGCACTATGACTTTGTGGGCTCCTATGAGAGGCTGGAGGCTGATGCAAATC

AGGTGCTGGAGTGGGTACGGGCACCACCTCACGTCCGATTTCCAGCTCGCCAGGCCTGGTACC

GGCCAGCCAGCCCCGAAAGCCTGCATTACCACTTGTGCAGTGCCCCCCGGGCCCTGCTGCAGG

ATGTGCTGCCTAAGTATATCCTGGACTTCTCCCTCTTTGCCTACCCACTGCCTAATGTCACCA

AGGAGGCGTGTCAGCAG**TGA**CCATGGGTGTGGGGCCAGCAGCTGGTGGGGACTGGTTTCAACG

CCAGCTTTCTGTGCTTCTGCCTGTCATTCGGAGAAACTCTGGCTCTGGGGCTTGGGGCTTCTC

AGGATCCTGGATGGCAGAGACTGCCCTCAGAAGTTCCTTGTCCAGGGTGGGCACCCACAGTGA

CTCAGAGGACAGGGCTAGGCAGGAGACCTGCTGCTCCTCATTGGGGGGATCTCTTGGGGGGCA

GACACCAGTTTGCCAATGAAGCAACACATCTGATCTAAAGACTGGCTCCAGACCCCGGGCTGC

CAGGATTATGCAGTCCACTTGGTCTACCTTAATTTAACCTGTGGCCAAACTCAGAGATGGTAC

CAGCCAGGGGCAAGCATGACCAGAGCCAGGGACCCTGTGGCTCTGATCCCCCATTTATCCACC

CCATGTGCCTCAGGACTAGAGTGAGCAATCATACCTTATAAATGACTTTTGTGCCTTTCTGCT

CCAGTCTCAAAATTTCCTACACCTGCCAGTTCTTTACATTTTTCCAAGGAAAGGAAAACGGAA

GCAGGGTTCTTGCCTGGTAGCTCCAGGACCCAGCTCTGCAGGCACCCAAAGACCCTCTGTGCC

CAGCCTCTTCCTTGAGTTCTCGGAACCTCCTCCCTAATTCTCCCTTCCTTCCCCACAAGGCCT

TTGAGGTTGTGACTGTGGCTGGTATATCTGGCTGCCATTTTTCTGATGCATTTATTTAAAATT

TGTACTTTTTGATAGAACCCTTGTAAGGGCTTTGTTTTCCTAATAGCTGACTTTTTAATAAAG

CAGTTTTATATAT

FIGURE 132

MLMFAVIVASSGLLLMIERGILAEMKPLPLHPPGREGTAWRGKAPKPGGLSLRAGDADLQVRQ

DVRNRTLRAVCGQPGMPRDPWDLPVGQRRTLLRHILVSDRYRFLYCYVPKVACSNWKRVMKVL

AGVLDSVDVRLKMDHRSDLVFLADLRPEEIRYRLQHYFKFLFVREPLERLLSAYRNKFGEIRE

YQQRYGAEIVRRYRAGAGPSPAGDDVTFPEFLRYLVDEDPERMNEHWMPVYHLCQPCAVHYDF

VGSYERLEADANQVLEWVRAPPHVRFPARQAWYRPASPESLHYHLCSAPRALLQDVLPKYILD

FSLFAYPLPNVTKEACQQ

Important features of the protein:

Signal peptide:

amino acids 1-23

N-glycosylation sites.

amino acids 67-71, 325-329

Tyrosine kinase phosphorylation sites.

amino acids 152-159, 183-183

N-myristoylation sites.

amino acids 89-95, 128-134

FIGURE 133

CGGCAGTTCTGGCCCCTGCAGCTGGAGGTACCCTGAGTTCTGAGGGTCGTAGTGCTGTTTCTG

GTATTCTCATCGCGGTCACCTCTACCGGTGTGGACAAGTAAAGTTTGAATCAGCTTCTCCATG

GCCTGGGCACCAGTTCCCGGCTGAGCCATTTTCCTTTTGGCTAAAAGTCCCCGCCCAGAGGCC

AATTCGTCGCGGCGGCGGTGGAGATCGCAGGTCGCTCAGGCTTGCAG**ATG**GGTCAAGGGTTGT

GGAGAGTGGTCAGAAACCAGCAGCTGCAACAAGAAGGCTACAGTGAGCAAGGCTACCTCACCA

GAGAGCAGAGCAGGAGAATGGATGCGAGCAACATTTCTAACACCAATCATCGTAAACAAGTCC

AAGGAGGCATTGACATATATCATCTTTTGAAGGCAAGGAAATCGAAAGAACAGGAAGGATTCA

TTAATTTGGAAATGTTGCCTCCTGAGCTAAGCTTTACCATCTTGTCCTACCTGAATGCAACTG

ACCTTTGCTTGGCTTCATGTGTTTGGCAGGACCTTGCGAATGATGAACTTCTCTGGCAAGGGT

TGTGCAAATCCACTTGGGGTCACTGTTCCATATACAATAAGAACCCACCTTTAGGATTTTCTT

TTAGAAAATTGTATATGCAGCTGGATGAAGGCAGCCTCACCTTTAATGCCAACCCAGATGAGG

GAGTGAACTACTTTATGTCCAAGGGTATCCTGGATGATTCGCCAAAGGAAATAGCAAAGTTTA

TCTTCTGTACAAGAACACTAAATTGGAAAAAACTGAGAATCTATCTTGATGAAAGGAGAGATG

TCTTGGATGACCTTGTAACATTGCATAATTTTAGAAATCAGTTCTTGCCAAATGCACTGAGAG

AATTTTTTCGTCATATCCATGCCCCTGAAGAGCGTGGAGAGTATCTTGAAACTCTTATAACAA

AGTTCTCACATAGATTCTGTGCTTGCAACCCTGATTTAATGCGAGAACTTGGCCTTAGTCCTG

ATGCTGTCTATGTACTGTGCTACTCTTTGATTCTACTTTCCATTGACCTCACTAGCCCTCATG

TGAAGAATAAAATGTCAAAAAGGGAATTTATTCGAAATACCCGTCGCGCTGCTCAAAATATTA

GTGAAGATTTTGTAGGGCATCTTTATGACAATATCTACCTTATTGGCCATGTGGCTGCA**TAA**A

AAGCACAATTGCTAGGACTTCAGTTTTTACTTCAGACTAAAGCTACCCAAGGACTTAGCAGAT

ATGGGGGTTACATCAGTGCTGGTCATTGTAGCCTGAGTATACAATCAAGCTTCAGTGTGCAAC

CTTTTTTTCTTTTGCCATTTTCTATTTTAGTAATTTCCTTGGGGAACTAAATAATTTTGCAGA

ATTTTTCCTAATTTTGTTTATCACGTTTTGCACAAAGCAGAGCCACTGTCTAACACAGCTGTT

AACGAATGATAAACTGACATTATACTCTAAAAGATGGTGTATTTGTGCATTAGATTTGCCTGA

AAAACTTTATCCATTTCCATTCTTTATACAAATACCATGTAATGTGTACATATTTAACTAAAG

AGATTTATAGTCATAATTATTTTATTGTAAAGATTTTAACTAAAGTTTTTCCTTTTCTCTC

FIGURE 134

MGQGLWRVVRNQQLQQEGYSEQGYLTREQSRRMDASNISNTNHRKQVQGGIDIYHLLKARKSK
EQEGFINLEMLPPELSFTILSYLNATDLCLASCVWQDLANDELLWQGLCKSTWGHCSIYNKNP
PLGFSFRKLYMQLDEGSLTFNANPDEGVNYFMSKGILDDSPKEIAKFIFCTRTLNWKKLRIYL
DERRDVLDDLVTLHNFRNQFLPNALREFFRHIHAPEERGEYLETLITKFSHRFCACNPDLMRE
LGLSPDAVYVLCYSLILLSIDLTSPHVKNKMSKREFIRNTRRAAQNISEDFVGHLYDNIYLIG
HVAA

Important features of the protein:
Transmembrane domain:
amino acids 253-272

N-glycosylation sites.
amino acids 37-41, 87-91, 298-302

N-myristoylation site.
amino acids 110-116

FIGURE 135

GGCACGAGGGAGCCTCCGTTAGGGGGTGGGAAAGGACTTTGCCATAGGTCGCTGAGGCCACCA
TCTGCTCTCTTACTGGCCAAGGGCGTAAAAAGATAGTCTTCCCATTAGCTAGAGAGCAAACCC
CAGAAAGCCTATTGGCTGCGCCGTCCGCGGGCCTTGGTCCGCTTTGAAGGCGGGCTGCGGCTG
CGAGAGGAGGGCGGGCGGGAGGCTAGCTGTTGTCGTGGTTGCTCGGAGGCACGTGTGCAGTCC
CGGAAGCGGCGAGGGGAAACTGCTCCGCGCGCGCCGCGGGAGGAGGAACCGCCCGGTCCTTTA
GGGTCCGGGCCCGGCCGGGCC**ATG**GATTCAATGCCTGAGCCCGCGTCCCGCTGTCTTCTGCTT
CTTCCCTTGCTGCTGCTGCTGCTGCTGCTGCCGGCCCCGGAGCTGGGCCCGAGCCAGGCC
GGAGCTGAGGAGAACGACTGGGTTCGCCTGCCCAGCAAATGCGAAGTGTGTAAATATGTTGCT
GTGGAGCTGAAGTCAGCCTTTGAGGAAACCGGCAAGACCAAGGAGGTGATTGGCACGGGCTAT
GGCATCCTGGACCAGAAGGCCTCTGGAGTCAAATACACCAAGTCGGACTTGCGGTTAATCGAA
GTCACTGAGACCATTTGCAAGAGGCTCCTGGATTATAGCCTGCACAAGGAGAGGACCGGCAGC
AATCGATTTGCCAAGGGCATGTCAGAGACCTTTGAGACATTACACAACCTGGTACACAAAGGG
GTCAAGGTGGTGATGGACATCCCCTATGAGCTGTGGAACGAGACTTCTGCAGAGGTGGCTGAC
CTCAAGAAGCAGTGTGATGTGCTGGTGGAAGAGTTTGAGGAGGTGATCGAGGACTGGTACAGG
AACCACCAGGAGGAAGACCTGACTGAATTCCTCTGCGCCAACCACGTGCTGAAGGGAAAAGAC
ACCAGTTGCCTGGCAGAGCAGTGGTCCGGCAAGAAGGGAGACACAGCTGCCCTGGGAGGGAAG
AAGTCCAAGAAGAAGAGCAGCAGGGCCAAGGCAGCAGGCGGCAGGAGTAGCAGCAGCAAACAA
AGGAAGGAGCTGGGTGGCCTTGAGGGAGACCCCAGCCCCGAGGAGGATGAGGGCATCCAGAAG
GCATCCCCTCTCACACACAGCCCCCCTGATGAGCTC**TGA**GCCCACCCAGCATCCTCTGTCCTG
AGACCCCTGATTTTGAAGCTGAGGAGTCAGGGGCATGGCTCTGGCAGGCCGGGATGGCCCCGC
AGCCTTCAGCCCCTCCTTGCCTTGGCTGTGCCCTCTTCTGCCAAGGAAAGACACAAGCCCCAG
GAAGAACTCAGAGCCGTCATGGGTAGCCCACGCCGTCCTTTCCCCTCCCCAAGTGTTTCTCTC
CTGACCCAGGGTTCAGGCAGGCCTTGTGGTTTCAGGACTGCAAGGACTCCAGTGTGAACTCAG
GAGGGGCAGGTGTCAGAACTGGGCACCAGGACTGGAGCCCCCTCCGGAGACCAAACTCACCAT
CCCTCAGTCCTCCCCAACAGGGTACTAGGACTGCAGCCCCCTGTAGCTCCTCTCTGCTTACCC
CTCCTGTGGACACCTTGCACTCTGCCTGGCCCTTCCCAGAGCCCAAAGAGTAAAAATGTTCTG
GTTCTGATTTCTGAAAAAAAAAAAAAAAAAAAAATTCCT

FIGURE 136

MDSMPEPASRCLLLLPLLLLLLLLLLPAPELGPSQAGAEENDWVRLPSKCEVCKYVAVELKSAF
EETGKTKEVIGTGYGILDQKASGVKYTKSDLRLIEVTETICKRLLDYSLHKERTGSNRFAKGM
SETFETLHNLVHKGVKVVMDIPYELWNETSAEVADLKKQCDVLVEEFEEVIEDWYRNHQEEDL
TEFLCANHVLKGKDTSCLAEQWSGKKGDTAALGGKKSKKKSSRAKAAGGRSSSSKQRKELGGL
EGDPSPEEDEGIQKASPLTHSPPDEL

Important features of the protein:
Signal peptide:
amino acids 1-26

N-glycosylation site.
amino acids 153-157

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 227-231, 228-232

Tyrosine kinase phosphorylation site.
amino acids 142-150

N-myristoylation sites.
amino acids 36-42, 74-80, 86-92, 125-131, 222-228, 237-243, 250-256, 263-269

Amidation sites.
amino acids 212-216, 222-226

ATP/GTP-binding site motif A (P-loop).
amino acids 62-70

FIGURE 137

```
CACGCCTCCCGCTGCCAGCCCGGCACCGGGATCTTAATCAGTCACTATGAAAACTCATTAGCT
CCACAGCAATGAGTCCTCCACTGCTGAAGCTTGGCGCTGTGCTTAGTACCATGGCAATGATCT
CAAACTGGATGTCCCAAACTCTCCCATCCTTGGTGGGACTGAACACCACGAGGCTGTCGACTC
CGGATACCTTAACTCAGATTAGTCCTAAAGAAGGGTGGCAGGTGTACAGCTCAGCTCAGGATC
CTGATGGGCGGTGCATTTGCACAGTTGTTGCTCCAGAACAAAACCTGTGTTCCCGGGATGCCA
AAAGCAGGCAACTTCGCCAACTACTGGAAAAGGTTCAGAACATGTCCCAGTCTATTGAAGTCT
TAAACTTGAGAACTCAGAGAGATTTCCAATATGTTTTAAAAATGGAAACCCAAATGAAAGGGC
TGAAGGCAAAATTTCGGCAGATTGAAGATGATCGAAAGACACTTATGACCAAGCATTTTCAGG
AGTTGAAAGAGAAATGGACGAGCTCCTGCCTTTGATCCCCGTGCTGGAACAGTACAAAACAG
ATGCTAAGTTAATCACCCAGTTCAAGGAGGAAATAAGGAATCTGTCTGCTGTCCTCACTGGTA
TTCAGGAGGAAATTGGTGCCTATGACTACGAGGAACTACACCAAAGAGTGCTGAGCTTGGAAA
CAAGACTTCGTGACTGCATGAAAAAGCTAACATGTGGCAAACTGATGAAAATCACAGGCCCAG
TTACAGTCAAGACATCTGGAACCCGATTTGGTGCTTGGATGACAGACCCTTTAGCATCTGAGA
AAAACAACAGAGTCTGGTACATGGACAGTTATACTAACAATAAAATTGTTCGTGAATACAAAT
CAATTGCAGACTTTGTCAGTGGGGCTGAATCAAGGACATACAACCTTCCTTTCAAGTGGGCAG
GAACTAACCATGTTGTCTACAATGGCTCACTCTATTTTAACAAGTATCAGAGTAATATCATCA
TCAAATACAGCTTTGATATGGGGAGAGTGCTTGCCCAACGAAGCCTGGAGTATGCTGGTTTTC
ATAATGTTTACCCCTACACATGGGGTGGATTCTCTGACATCGACCTAATGGCTGATGAAATCG
GGCTGTGGGCTGTGTATGCAACTAACCAGAATGCAGGCAATATTGTCATCAGCCAACTTAACC
AAGATACCTTGGAGGTGATGAAGAGCTGGAGCACTGGCTACCCCAAGAGAAGTGCAGGGGAAT
CTTTCATGATCTGTGGGACACTGTATGTCACCAACTCCCACTTAACTGGAGCCAAGGTGTATT
ATTCCTATTCCACCAAAACCTCCACATATGAGTACACAGACATTCCCTTCCATAACCAATACT
TTCACATATCCATGCTTGACTACAATGCAAGAGATCGAGCTCTCTATGCCTGGAACAATGGCC
ACCAGGTGCTGTTCAATGTCACCCTTTTCCATATCATCAAGACAGAGGATGACACATAGGCAA
ATGTGACATGTTTTCATTGATTTAAACAGTGTGATTTGTGATAAACTCTATAAGACCCCTTCC
GTTTTTTTCTTCACTATTATTTTTCATCATTTCTCCAAAGCAAAGCATTTTTATTGTAAAGTT
GGTGTTTCAAAAACATAGCTGAGCTTGTCTAACTTACCATGTTGGAAACACATCTTAACTTCT
AAATTTACAAGGCCTATCATGTCCTTGTCATGAAAAGCACTAAAAAAAAAAAAAGAGTTAAGT
GGCTAAAGTCATAGTTTTGCAAGAGATTAATGATCTGCCTTATATTAGAGTCAGAGACTAATG
GTGGCTTAAATGCACGAATGTCTTTTTTTTAAAACTGTCATTTTTACTGTCTTTTGCTCCA
TCTCAGGAAATATTTTGGTAGGAATTAGGAGAACAAAAGCACTTTTATCCCATTTATTTCTT
TAAAAAATGTAAGGATTTCATTTATATTGAAAATAATATTAATCATTTTGCTGTTAACACAA
TTCTCTGATGCGGTGCTGTACAGTCATTTTTAAATCTCTTGCTAACATTTTATTGGCAGTATG
TATTTCTACCATTGTAACCACCATTGTGCTATTGTATCTCTTCACTTCTGTGAAAGTAATATT
TTTTATAAAANACACTGNAATTTTAAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAA
```

FIGURE 138

MSPPLLKLGAVLSTMAMISNWMSQTLPSLVGLNTTRLSTPDTLTQISPKEGWQVYSSAQDPDG
RCICTVVAPEQNLCSRDAKSRQLRQLLEKVQNMSQSIEVLNLRTQRDFQYVLKMETQMKGLKA
KFRQIEDDRKTLMTKHFQELKEKMDELLPLIPVLEQYKTDAKLITQFKEEIRNLSAVLTGIQE
EIGAYDYEELHQRVLSLETRLRDCMKKLTCGKLMKITGPVTVKTSGTRFGAWMTDPLASEKNN
RVWYMDSYTNNKIVREYKSIADFVSGAESRTYNLPFKWAGTNHVVYNGSLYFNKYQSNIIIKY
SFDMGRVLAQRSLEYAGFHNVYPYTWGGFSDIDLMADEIGLWAVYATNQNAGNIVISQLNQDT
LEVMKSWSTGYPKRSAGESFMICGTLYVTNSHLTGAKVYYSYSTKTSTYEYTDIPFHNQYFHI
SMLDYNARDRALYAWNNGHQVLFNVTLFHIIKTEDDT

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 33-37, 95-99, 179-183, 299-303, 465-469

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 215-219

Tyrosine kinase phosphorylation site.

amino acids 106-114

N-myristoylation sites.

amino acids 9-15, 31-37, 235-241, 239-245

FIGURE 139

GAAGCAGTGCAGAGAGGAGAGCGGAGCGGAGCTGCCGCTGAGCAAAGGCCTTCACC**ATG**GCCG

AGTCCCCCGGCTGCTGCTCCGTCTGGGCCCGCTGCCTCCACTGCCTGTATAGCTGCCACTGGA

GGAAATGCCCCAGAGAGAGGATGCAAACCAGCAAGTGCGACTGTATCTGGTTTGGCCTGCTCT

TCCTCACCTTCCTCCTTTCCCTGAGCTGGCTGTACATCGGGCTCGTCCTTCTCAATGACCTGC

ACAACTTCAATGAATTCCTCTTCCGCCGCTGGGGACACTGGATGGACTGGTCCCTGGCATTCC

TGCTGGTCATCTCTCTACTGGTCACATATGCATCCTTGCTATTGGTCCTGGCCCTGCTCCTGC

GGCTTTGTAGACAGCCCCTGCATCTGCACAGCCTCCACAAGGTGCTGCTGCTCCTCATTATGC

TGCTTGTGGCGGCTGGCCTTGTGGGACTGGACATCCAATGGCAGCAGGAGTGGCATAGCTTGC

GTGTGTCACTGCAGGCCACAGCCCCATTCCTTCATATTGGAGCAGCCGCTGGAATTGCCCTCC

TGGCCTGGCCTGTGGCTGATACCTTCTACCGTATCCACCGAAGAGGTCCCAAGATTCTGCTAC

TGCTCCTATTTTTTGGAGTTGTCCTGGTCATCTACTTGGCCCCCCTATGCATCTCCTCACCCT

GCATCATGGAACCCAGAGACTTACCACCCAAGCCTGGGCTGGTGGGACACCGAGGGGCCCCCA

TGCTGGCTCCCGAGAACACCCTGATGTCCTTGCGGAAGACAGCTGAATGCGGAGCTACTGTGT

TTGAGACTGATGTGATGGTCAGCTCCGATGGGGTCCCCTTCCTCATGCATGATGAGCACCTCA

GCAGGACCACGAATGTAGCCTCTGTATTCCCAACCCGAATCACAGCCCACAGCAGTGACTTCT

CCTGGACTGAACTGAAGAGACTCAATGCTGGATCCTGGTTCCTAGAGAGGCGACCCTTCTGGG

GGGCCAAACCGCTGGCAGGCCCTGATCAGAAAGAGGCTGAGAGTCAGACGGTACCAGCATTAG

AAGAGCTATTGGAGGAAGCTGCAGCCCTCAACCTTTCCATCATGTTCGACTTGCGCCGACCCC

CACAGAACCACACATACTATGACACTTTTGTGATCCAGACATTGGAGACTGTGCTGAATGCAA

GGGTGCCCCAAGCCATGGTCTTTTGGCTACCAGATGAAGATCGGGCTAATGTCCAACGACGGG

CACCTGGAATGCGCCAGATATATGGACGTCAGGGAGGCAACAGAACGGAGAGGCCCCAGTTTC

TTAACCTCCCCTATCAAGATCTGCCACTATTGGATATCAAGGCATTGCATAAGGATAATGTCT

CGGTGAACCTATTTGTAGTGAACAAGCCCTGGCTCTTCTCTCTGCTTTGGTGTGCAGGGGTGG

ATTCGGTCACCACCAACGACTGCCAGCTGCTGCAGCAGATGCGTTACCCTATCTGGCTTATTA

CCCCTCAAACCTACCTAATCATATGGGTCATTACCAATTGTGTTTCCACCATGCTGCTTTTGT

GGACCTTCCTCCTCCAAAGGAGATTTGTTAAGAAGAGAGGGAAAACTGGCTTAGAAACAGCAG

TGCTGCTGACAAGGATCAACAATTTCATGATGGAG**TGA**ATGCCCTGCCCTGCTTCCCCACCCA

AGCCAGTCTACATTGCCCAAACAGCAAGGGTTGGAGAGTGGCTTAAGTGGAATGCTTCAGGGG

TGGTGGGTTGCAAGTGGGGGGAGCTTTGCCAACAGGAGGTTTTGAACCATGAGGGCCCTCTGC

CCAGGTGATGGGCATTCCCTAAGCTGCTATGGAATCTGCTCCCTTTGGGGTTTTGACCTGAGA

TGTTTGGGAAGAGAGTGAGTAATGAGAAGTTTCTCCTCAAATGAAACTAGAACAGAGGAAGTA

AAAGGGAGATTGCTCGGA

FIGURE 140

MAESPGCCSVWARCLHCLYSCHWRKCPRERMQTSKCDCIWFGLLFLTFLLSLSWLYIGLVLLN
DLHNFNEFLFRRWGHWMDWSLAFLLVISLLVTYASLLLVLALLLRLCRQPLHLHSLHKVLLLL
IMLLVAAGLVGLDIQWQQEWHSLRVSLQATAPFLHIGAAAGIALLAWPVADTFYRIHRRGPKI
LLLLLFFGVVLVIYLAPLCISSPCIMEPRDLPPKPGLVGHRGAPMLAPENTLMSLRKTAECGA
TVFETDVMVSSDGVPFLMHDEHLSRTTNVASVFPTRITAHSSDFSWTELKRLNAGSWFLERRP
FWGAKPLAGPDQKEAESQTVPALEELLEEAAALNLSIMFDLRRPPQNHTYYDTFVIQTLETVL
NARVPQAMVFWLPDEDRANVQRRAPGMRQIYGRQGGNRTERPQFLNLPYQDLPLLDIKALHKD
NVSVNLFVVNKPWLFSLLWCAGVDSVTTNDCQLLQQMRYPIWLITPQTYLIIWVITNCVSTML
LLWTFLLQRRFVKKRGKTGLETAVLLTRINNFMME

Important features of the protein:
Transmembrane domains:
amino acids 38-60, 83-107, 122-138, 156-173, 189-210, 484-506

N-glycosylation sites.
amino acids 349-353, 362-366, 415-419, 442-446

N-myristoylation sites.
amino acids 163-169, 413-419, 523-529

Leucine zipper pattern.
amino acids 93-115, 109-131

Glutamine amidotransferases class-II active site.
amino acids 1-13

FIGURE 141

GCCGCCGGCCCGGGCTGGAGCCGAGCGCAGCAGCCACCGCCGCCGCCGCGCCAGAAGTTTGGGTTGAACCGGAGC
TGCCGGGGAGGAAACTTTTTTCTTTTTTCCCCCTCCCTCCCGGGAGGAGGAGGAGGAGGAGGGGAAGCTGCCG
CCGGCGCCAAGGCTCGTGGGCTCGGGGTCGGCGCGGCCCGCAGAAGGGGCGGGGGCCTCGCCCCGCGAGGGGAGG
CGCGCCCCGGGGGCCCCGAGAGGGGCGGTGAGGACCGCGGGCTGCTGGTGCGGCGGCGGCGGCGCGTGTGCCCCG
CGCAGGGGAGGGCGCCCGCCCCGCTCCCGGCCCGGCTGCGAGGAGGAGGCGGCGGCGGCGCAGGAG**ATG**TACTT
GGTGGCGGGGGACAGGGGGTTGGCCGGCTGCGGGCACCTCCTGGTCTCGCTGCTGGGGCTGCTGCTGCTGGC
GCGCTCCGGCACCCGGGCGCTGGTCTGCCTGCCCTGTGACGAGTCCAAGTGCGAGGAGCCCAGGAACTGCCCGGG
GAGCATCGTGCAGGGCGTCTGCGGCTGCTGCTACACGTGCGCCAGCCAGAGGAACGAGAGCTGCGGCGGCACCTT
CGGGATTTACGGAACCTGCGACCGGGGGCTGCGTTGTGTCATCCGCCCCCCGCTCAATGGCGACTCCCTCACCGA
GTACGAAGCGGGCGTTTGCGAAGATGAGAACTGGACTGATGACCAACTGCTTGGTTTTAAACCATGCAATGAAAA
CCTTATTGCTGGCTGCAATATAATCAATGGGAAATGTGAATGTAACACCATTCGAACCTGCAGCAATCCCTTTGA
GTTTCCAAGTCAGGATATGTGCCTTTCAGCTTTAAAGAGAATTGAAGAAGAGAAGCCAGATTGCTCCAAGGCCCG
CTGTGAAGTCCAGTTCTCTCCACGTTGTCCTGAAGATTCTGTTCTGATCGAGGGTTATGCTCCTCCTGGGGAGTG
CTGTCCCTTACCCAGCCGCTGCGTGTGCAACCCCGCAGGCTGTCTGCGCAAAGTCTGCCAGCCGGGAAACCTGAA
CATACTAGTGTCAAAAGCCTCAGGGAAGCCGGGAGAGTGCTGTGACCTCTATGAGTGCAAACCAGTTTTCGGCGT
GGACTGCAGGACTGTGGAATGCCCTCCTGTTCAGCAGACCGCGTGTCCCCCGGACAGCTATGAAACTCAAGTCAG
ACTAACTGCAGATGGTTGCTGTACTTTGCCAACAAGATGCGAGTGTCTCTCTGGCTTATGTGGTTTCCCCGTGTG
TGAGGTGGGATCCACTCCCCGCATAGTCTCTCGTGGCGATGGGACACCTGGAAAGTGCTGTGATGTCTTTGAATG
TGTTAATGATACAAAGCCAGCCTGCGTATTTAACAATGTGGAATATTATGATGGAGACATGTTTCGAATGGACAA
CTGTCGGTTCTGTCGATGCCAAGGGGGCGTTGCCATCTGCTTCACTGCCCAGTGTGGTGAGATAAACTGCGAGAG
GTACTACGTGCCCGAAGGAGAGTGCTGCCCAGTGTGTGAAGATCCAGTGTATCCTTTTAATAATCCCGCTGGCTG
CTATGCCAATGGCCTGATCCTTGCCCACGGAGACCGGTGGCGGGAAGACGACTGCACATTCTGCCAGTGCGTCAA
CGGTGAACGCCACTGCGTTGCGACCGTCTGCGGACAGACCTGCACAAACCCTGTGAAAGTGCCTGGGGAGTGTTG
CCCTGTGTGCGAAGAACCAACCATCATCACAGTTGATCCACCTGCATGTGGGGAGTTATCAAACTGCACTCTGAC
AGGGAAGGACTGCATTAATGGTTTCAAACGCGATCACAATGGTTGTCGGACCTGTCAGTGCATAAACACCGAGGA
ACTATGTTCAGAACGTAAACAAGGCTGCACCTTGAACTGTCCCTTCGGTTTCCTTACTGATGCCCAAAACTGTGA
GATCTGTGAGTGCCGCCCAAGGCCCAAGAAGTGCAGACCCATAATCTGTGACAAGTATTGTCCACTTGGATTGCT
GAAGAATAAGCACGGCTGTGACATCTGTCGCTGTAAGAAATGTCCAGAGCTCTCATGCAGTAAGATCTGCCCCTT
GGGTTTCCAGCAGGACAGTCACGGCTGTCTTATCTGCAAGTGCAGAGAGGCCTCTGCTTCAGCTGGGCCACCCAT
CCTGTCGGGCACTTGTCTCACCGTGGATGGTCATCATCATAAAAATGAGGAGAGCTGGCACGATGGGTGCCGGGA
ATGCTACTGTCTCAATGGACGGGAAATGTGTGCCCTGATCACCTGCCCGGTGCCTGCCTGTGGCAACCCCACCAT
TCACCCTGGACAGTGCTGCCCATCATGTGCAGATGACTTTGTGGTGCAGAAGCCAGAGCTCAGTACTCCCTCCAT
TTGCCACGCCCCTGGAGGAGAATACTTTGTGGAAGGAGAAACGTGGAACATTGACTCCTGTACTCAGTGCACCTG
CCACAGCGGACGGGTGCTGTGTGAGACAGAGGTGTGCCCACCGCTGCTCTGCCAGAACCCCTCACGCACCCAGGA
TTCCTGCTGCCCACAGTGTACAGATCAACCTTTTCGGCCTTCCTTGTCCCGCAATAACAGCGTACCTAATTACTG
CAAAAATGATGAAGGGGATATATTCCTGGCAGCTGAGTCCTGGAAGCCTGACGTTTGTACCAGCTGCATCTGCAT
TGATAGCGTAATTAGCTGTTTCTCTGAGTCCTGCCCTTCTGTATCCTGTGAAAGACCTGTCTTGAGAAAAGGCCA
GTGTTGTCCCTACTGCATAGAAGACACAATTCCAAAGAAGGTGGTGTGCCACTTCAGTGGGAAGGCCTATGCCGA
CGAGGAGCGGTGGGACCTTGACAGCTGCACCCACTGCTACTGCCTGCAGGGCCAGACCCTCTGCTCGACCGTCAG
CTGCCCCCCTCTGCCCTGTGTTGAGCCCATCAACGTGGAAGGAAGTTGCTGCCCAATGTGTCCAGAAATGTATGT
CCCAGAACCAACCAATATACCCATTGAGAAGACAAACCATCGAGGAGAGGTTGACCTGGAGGTTCCCCTGTGGCC
CACGCCTAGTGAAAATGATATCGTCCATCTCCCTAGAGATATGGGTCACCTCCAGGTAGATTACAGAGATAACAG
GCTGCACCCAAGTGAAGATTCTTCACTGGACTCCATTGCCTCAGTTGTGGTTCCCATAATTATATGCCTCTCTAT
TATAATAGCATTCCTATTCATCAATCAGAAGAAACAGTGGATACCACTGCTTTGCTGGTATCGAACACCAACTAA
GCCTTCTTCCTTAAATAATCAGCTAGTATCTGTGGACTGCAAGAAAGGAACCAGAGTCCAGGTGGACAGTTCCCA
GAGAATGCTAAGAATTGCAGAACCAGATGCAAGATTCAGTGGCTTCTACAGCATGCAAAAACAGAACCATCTACA
GGCAGACAATTTCTACCAAACAGTG**TGA**AGAAAGGCAACTAGGATGAGGTTTCAAAAGACGGAAGACGACTAAAT
CTGCTCTAAAAAGTAAACTAGAATTTGTGCACTTGCTTAGTGGATTGTATTGGATTGTGACTTGATGTACAGCGC
TAAGACCTTACTGGGATGGGCTCTGTCTACAGCAATGTGCAGAACAAGCATTCCCACTTTTCCTCAAAAAA

FIGURE 142

MYLVAGDRGLAGCGHLLVSLLGLLLLLARSGTRALVCLPCDESKCEEPRNCPGSIVQGVCGCC
YTCASQRNESCGGTFGIYGTCDRGLRCVIRPPLNGDSLTEYEAGVCEDENWTDDQLLGFKPCN
ENLIAGCNIINGKCECNTIRTCSNPFEFPSQDMCLSALKRIEEEKPDCSKARCEVQFSPRCPE
DSVLIEGYAPPGECCPLPSRCVCNPAGCLRKVCQPGNLNILVSKASGKPGECCDLYECKPVFG
VDCRTVECPPVQQTACPPDSYETQVRLTADGCCTLPTRCECLSGLCGFPVCEVGSTPRIVSRG
DGTPGKCCDVFECVNDTKPACVFNNVEYYDGDMFRMDNCRFCRCQGGVAICFTAQCGEINCER
YYVPEGECCPVCEDPVYPFNNPAGCYANGLILAHGDRWREDDCTFCQCVNGERHCVATVCGQT
CTNPVKVPGECCPVCEEPTIITVDPPACGELSNCTLTGKDCINGFKRDHNGCRTCQCINTEEL
CSERKQGCTLNCPFGFLTDAQNCEICECRPRPKKCRPIICDKYCPLGLLKNKHGCDICRCKKC
PELSCSKICPLGFQQDSHGCLICKCREASASAGPPILSGTCLTVDGHHHKNEESWHDGCRECY
CLNGREMCALITCPVPACGNPTIHPGQCCPSCADDFVVQKPELSTPSICHAPGGEYFVEGETW
NIDSCTQCTCHSGRVLCETEVCPPLLCQNPSRTQDSCCPQCTDQPFRPSLSRNNSVPNYCKND
EGDIFLAAESWKPDVCTSCICIDSVISCFSESCPSVSCERPVLRKGQCCPYCIEDTIPKKVVC
HFSGKAYADEERWDLDSCTHCYCLQGQTLCSTVSCPPLPCVEPINVEGSCCPMCPEMYVPEPT
NIPIEKTNHRGEVDLEVPLWPTPSENDIVHLPRDMGHLQVDYRDNRLHPSEDSSLDSIASVVV
PIIICLSIIIAFLFINQKKQWIPLLCWYRTPTKPSSLNNQLVSVDCKKGTRVQVDSSQRMLRI
AEPDARFSGFYSMQKQNHLQADNFYQTV

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 940-962

N-glycosylation sites.

amino acids 71-75, 113-117, 330-334, 474-478, 746-750

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 992-996

N-myristoylation site.

amino acids 9-15, 58-64, 61-67, 75-81, 79-85, 362-368, 402-408, 407-413, 439-445, 492-498, 511-517, 551-557, 558-564, 586-592, 606-612, 625-631, 845-851

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 52-63, 844-855

Cell attachment sequence.

amino acids 314-317

Leucine zipper pattern.

amino acids 3-25

Eukaryotic thiol (cysteine) proteases cysteine active site.

amino acids 57-69

VWFC domain proteins.

amino acids 448-456, 382-390

C-terminal cystine knot proteins

amino acids 60-86

FIGURE 143

GACGTCTGGCCGGCTCCCGGCGAAGGGCAGCGGAGGAGCGGCCCAGAGCGCGCAGCTAGGGCA

CTGGCGAAACCCCGGGACAGTCCCTCTCCGTGCGGGGGCGGCGCAGAGCAGTCCCATCCCCGG

GGTCCCGGGCGCGGCTGACTGCCGGCTGGTTCCCTGCGCGCAGTAGCTCCCCGAGCCGGGCTG

CACCGGAGGCGGCGAG**ATG**GTCGCGCGCGTCGGCCTCCTGCTGCGCGCCCTGCAGCTGCTACT

GTGGGGCCACCTGGACGCCCAGCCCGCGGAGCGCGGAGGCCAGGAGCTGCGCAAGGAGGCGGA

GGCATTCCTAGAGAAGTACGGATACCTCAATGAACAGGTCCCCAAAGCTCCCACCTCCACTCG

ATTCAGCGATGCCATCAGAGCGTTTCAGTGGGTGTCCCAGCTACCTGTCAGCGGCGTGTTGGA

CCGCGCCACCCTGCGCCAGATGACTCGTCCCCGCTGCGGGGTTACAGATACCAACAGTTATGC

GGCCTGGGCTGAGAGGATCAGTGACTTGTTTGCTAGACACCGGACCAAAATGAGGCGTAAGAA

ACGCTTTGCAAAGCAAGGTAACAAATGGTACAAGCAGCACCTCTCCTACCGCCTGGTGAACTG

GCCTGAGCATCTGCCGGAGCCGGCAGTTCGGGGCGCCGTGCGCGCCGCCTTCCAGTTGTGGAG

CAACGTCTCAGCGCTGGAGTTCTGGGAGGCCCCAGCCACAGGCCCCGCTGACATCCGGCTCAC

CTTCTTCCAAGGGGACCACAACGATGGGCTGGGCAATGCCTTTGATGGCCCAGGGGGCGCCCT

GGCGCACGCCTTCCTGCCCCGCCGCGGCGAAGCGCACTTCGACCAAGATGAGCGCTGGTCCCT

GAGCCGCCGCCGCGGGCGCAACCTGTTCGTGGTGCTGGCGCACGAGATCGGTCACACGCTTGG

CCTCACCCACTCGCCCGCGCCGCGCGCGCTCATGGCGCCCTACTACAAGAGGCTGGGCCGCGA

CGCGCTGCTCAGCTGGGACGACGTGCTGGCCGTGCAGAGCCTGTATGGGAAGCCCCTAGGGGG

CTCAGTGGCCGTCCAGCTCCCAGGAAAGCTGTTCACTGACTTTGAGACCTGGGACTCCTACAG

CCCCCAAGGAAGGCGCCCTGAAACGCAGGGCCCTAAATACTGCCACTCTTCCTTCGATGCCAT

CACTGTAGACAGGCAACAGCAACTGTACATTTTTAAAGGGAGCCATTTCTGGGAGGTGGCAGC

TGATGGCAACGTCTCAGAGCCCCGTCCACTGCAGGAAAGATGGGTCGGGCTGCCCCCCAACAT

TGAGGCTGCGGCAGTGTCATTGAATGATGGAGATTTCTACTTCTTCAAAGGGGGTCGATGCTG

GAGGTTCCGGGGCCCCAAGCCAGTGTGGGGTCTCCCACAGCTGTGCCGGGCAGGGGGCCTGCC

CCGCCATCCTGACGCCGCCCTCTTCTTCCCTCCTCTGCGCCGCCTCATCCTCTTCAAGGGTGC

CCGCTACTACGTGCTGGCCCGAGGGGGACTGCAAGTGGAGCCCTACTACCCCGAAGTCTGCA

GGACTGGGGAGGCATCCCTGAGGAGGTCAGCGGCGCCCTGCCGAGGCCCGATGGCTCCATCAT

CTTCTTCCGAGATGACCGCTACTGGCGCCTCGACCAGGCCAAACTGCAGGCAACCACCTCGGG

CCGCTGGGCCACCGAGCTGCCCTGGATGGGCTGCTGGCATGCCAACTCGGGGAGCGCCCTGTT

C**TGA**AGGCACCTCCTCACCTCAGAAACTGGTGGTGCTCTCAGGGCAAAATCATGTTCCCCACC

CCCGGGGCAGAACCCCTCTTAGAAGCCTCTGAGTCCCTCTGCAGAAGACCGGGCAGCAAAGCC

TCCATCTGGAAGTCTGTCTGCCTTTGTTCCTTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 144

MVARVGLLLRALQLLLWGHLDAQPAERGGQELRKEAEAFLEKYGYLNEQVPKAPTSTRFSDAI
RAFQWVSQLPVSGVLDRATLRQMTRPRCGVTDTNSYAAWAERISDLFARHRTKMRRKKRFAKQ
GNKWYKQHLSYRLVNWPEHLPEPAVRGAVRAAFQLWSNVSALEFWEAPATGPADIRLTFFQGD
HNDGLGNAFDGPGGALAHAFLPRRGEAHFDQDERWSLSRRGRNLFVVLAHEIGHTLGLTHSP
APRALMAPYYKRLGRDALLSWDDVLAVQSLYGKPLGGSVAVQLPGKLFTDFETWDSYSPQGRR
PETQGPKYCHSSFDAITVDRQQQLYIFKGSHFWEVAADGNVSEPRPLQERWVGLPPNIEAAAV
SLNDGDFYFFKGGRCWRFRGPKPVWGLPQLCRAGGLPRHPDAALFFPPLRRLILFKGARYYVL
ARGGLQVEPYYPRSLQDWGGIPEEVSGALPRPDGSIIFFRDDRYWRLDQAKLQATTSGRWATE
LPWMGCWHANSGSALF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 164-168, 355-359

N-myristoylation sites.

amino acids 92-98, 153-159, 193-199, 202-208, 288-294, 368-374,
509-515

Amidation site.

amino acids 312-316

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 237-247

Matrixins cysteine switch

amino acids 231-262, 271-284

Hemopexin domain protein

amino acids 66-108, 231-262

FIGURE 145

```
GCCGGCTAGGGCGCCGGAGCCGCACGCAGCCGCGGGGCTCCGAGAGGCGCGCACTGGGGCTGGGACTGCGCGGCG
CCGCCGCTGCGAGCGCCACTGAGCGGTCGCGCAACTTCGGAGGCACAGCGCCGGAGCCAGGCGAGCGCTCAGAGA
CCCGGAGCCAGAGGGGCGCGCCGGAGCCTCGTTCGAGAGCCGGCGCCAGGCACCCACCGCGCTCCGAGTGCCAGG
CGGCCCTCCGCGCAGCGTGGCTTCCGCTGCCCCCACGGAAGGCACGGGCTGGCGCTGCCGGGCGCCGGGGAGGAC
GGCGAGGAGGAGGCGGCGGCGGCGGAGACGGCGGCGGCGAGACTGGGGCCAGGGAGACAGCCCTGGGGGAGAGGC
GCCCGAACCAGGCCGCGGGAGC**ATG**GGGGCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTG
CTGGGACCCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCGTCAGCGCC
AGCAGAGCCGCTGCCCTACTTCCTGCAGGAGCCACAGGACGCCTACATTGTGAAGAACAAGCCTGTGGAGCTCCG
CTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAGTGCAACGGCGAGTGGGTCAGCCAGAACGACCACGTCAC
ACAGGAAGGCCTGGATGAGGCCACCGGCCTGCGGGTGCGCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGA
GGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCAGGCACCACCAAGAGTCGCCG
AGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTGGACCA
TGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGT
CATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCCCGCCTGTC
GGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGT
CTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAA
GCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGGGGCCTTCTGCGAGGGCCAGGCATTCCAGAAGAC
CGCCTGCACCACCATCTGCCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTG
TGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGGACGCTGCT
CGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGCCACCTGCT
GGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTGGTCGTGGCAATCCTCATGGC
GGTGGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTGAC
TGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCC
TGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCAT
GACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGG
GCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGCCGCCTGGCACATACCCTAGCGATTTCGCCCGGGA
CACCCACTTCCTGCACCTGCGCAGCGCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAG
CAGCGTCAGCGGCACCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTGGTGCC
CAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACCCTCCCGCT
TTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGT
CATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCA
GGGCACTGGGAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAGCCCAG
GGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCCTATTCCCGCTCAGCAGTCAA
GCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGA
GGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAA
ACCGCTAATGTTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGGAGGAG
CAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCAC
TTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGA
GGGCCAGATATTCCAGCTGCATACCACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCC
TGGCAGCACTGTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAA
CAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCT
GAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGG
GGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGG
GGACTGC**TGA**GCCTCCTGGGACAGCGGGCTGGCAGGGACTGGCAGGAGGCAGGTGCAGGGAGGCCTGGGGCAGCC
TCCTGATGGGGATGTTTGGCCTCTGCTTCCTCCCAGTTCACAGCCAGAGTTGCCTCTCCTCCTCCTCTTCCCCAA
CCCCCAGACCATGACCAGCCTTAGAAAATCCATGTACTCTGTTGTTAGAGGGCCCAGAGTTCCTTCTCCACCCCC
GCTCTCTCTCTTGGCCTGAGATCTCTGTGCAGGAACCAAGATGGGGCTGAAGCCTCTGGAGGCAGTTGGTTGG
GGGCGGGCAGGCAGGAGGCCCTCCCTCCACCCCCCCACCCTCAGCCCGGCAACTTCTGGGTTCCGTGGGTTTTAG
TTCCGTTCTTCGTTTTCTTCCTCCGTTATTGATTTCTCCTTTCTCCCTAAGCCCCCTTCTGCTTCCACGCCCTTT
TCCTCTTTGAAGAGTCAAGTACAATTCAGACAAACTGCTTTCTCCTGTCCAAAAGCAAAAAGGCAAAGGAAAGAA
AGAAAGCTTCAGACCGCTAGTAAGGCTCAAAGAAGAAGAAAAACACCAAAACCACAAGGGAAAAGAAAAACCCAG
TTTCTTAGGAAACGCAAACGATTTATTATCCAGATTATTTGGATAAGTCCTTTTTAAAA
```

FIGURE 146

```
MGARSGARGALLLALLLCWDPRLSQAGTDSGSEVLPDSFPSAPAEPLPYFLQEPQDAYIVKNK
PVELRCRAFPATQIYFKCNGEWVSQNDHVTQEGLDEATGLRVREVQIEVSRQQVEELFGLEDY
WCQCVAWSSAGTTKSRRAYVRIAYLRKNFDQEPLGKEVPLDHEVLLQCRPPEGVPVAEVEWLK
NEDVIDPTQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW
AEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQAFQKTACTTICPVDGAWTEWSKWSACS
TECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCMQNKKTLSDPNSHLLEASGDAALYA
GLVVAIFVVVAILMAVGVVVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPSNPQLLHPSV
PPDLTASAGIYRGPVYALQDSTDKIPMTNSPLLDPLPSLKVKVYSSSTTGSGPGLADGADLLG
VLPPGTYPSDFARDTHFLHLRSASLGSQQLLGLPRDPGSSVSGTFGCLGGRLSIPGTGVSLLV
PNGAIPQGKFYEMYLLINKAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTMPHCAEVSA
RDWIFQLKTQAHQGHWEEVVTLDEETLNTPCYCQLEPRACHILLDQLGTYVFTGESYSRSAVK
RLQLAVFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVEEPKPLMFKDSYHNLRL
SLHDLPHAHWRSKLLAKYQEIPFYHIWSGSQKALHCTFTLERHSLASTELTCKICVRQVEGEG
QIFQLHTTLAETPAGSLDTLCSAPGSTVTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRM
LAQKLSMDRYLNYFATKASPTGVILDLWEALQQDDGDLNSLASALEEMGKSEMLVAVATDGDC
```

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 374-395

N-glycosylation sites.

amino acids 222-225, 347-350

Glycosaminoglycan attachment site.

amino acids 492-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 233-236, 234-237

Casein kinase II phosphorylation sites.

amino acids 30-33, 87-90, 251-254, 341-344, 359-362, 629-632, 651-654, 706-709, 757-760, 827-830, 925-928, 941-944

Tyrosine kinase phosphorylation sites.

amino acids 216-223, 773-780

N-myristoylation sites.

amino acids 2-7, 6-11, 27-32, 96-101, 137-142, 179-184, 247-252, 281-286, 334-339, 379-384, 491-496, 495-500, 509-514, 542-547, 547-552, 550-555, 553-558, 560-565, 611-616, 785-790, 834-839, 844-849

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 541-551

ATP/GTP-binding site motif A (P-loop).

amino acids 926-933

Growth factor and cytokines receptors family signature 2.

amino acids 306-312

FIGURE 147

```
GAGAGGGACAGAGGCTGGAGAAGGATGTATGGCCTGCCCTGGGCTTGTCTGTTCCCTCCTGAGCCTGAGCCCCTT
ACCTTCCTGACCCCATGAAGCACACACTGGCTCTGCTGGCTCCCCTGCTGGGCCTGGGCCTGGGGCTGGCCCTGA
GTCAGCTGGCTGCAGGGGCCACAGACTGCAAGTTCCTTGGCCCGGCAGAGCACCTGACATTCACCCCAGCAGCCA
GGGCCCGGTGGCTGGCCCCTCGAGTTCGTGCGCCAGGACTCCTGGACTCCCTCTATGGCACCGTGCGCCGCTTCC
TCTCGGTGGTGCAGCTCAATCCTTTCCCTTCAGAGTTGGTAAAGGCCCTACTGAATGAGCTGGCCTCCGTGAAGG
TGAATGAGGTGGTGCGGTACGAGGCGGGCTACGTGGTATGCGCTGTGATCGCGGGCCTCTACCTGCTGCTGGTGC
CCACTGCCGGGCTTTGCTTCTGCTGCTGCCGCTGCCACCGGCGCTGCGGGGGACGAGTGAAGACAGAGCACAAGG
CGCTGGCCTGTGAGCGCGCGGCCCTCATGGTCTTCCTGCTGCTGACCACCCTCTTGCTGCTGATTGGTGTGGTCT
GTGCCTTTGTCACCAACCAGCGCACGCATGAACAGATGGGCCCCAGCATCGAGGCCATGCCTGAGACCCTGCTCA
GCCTCTGGGCCTGGTCTCTGATGTCCCCCAAGAGCTGCAGGCCGTGGCACAGCAATTCTCCCTGCCCCAGGAGC
AAGTCTCAGAGGAGCTGGATGGTGTTGGTGTGAGCATTGGGAGCGCGATCCACACTCAGCTCAGGAGCTCCGTGT
ACCCCTTGCTGGCGGCCGTGGGCAGTTTGGGCCAGGTCCTGCAGGTCTCCGTGCACCACCTGCAAACCTTGAATG
CTACAGTGGTAGAGCTGCAGGCCGGGCAGCAGGACCTGGAGGCCAGCCATCCGGGAACACCGGGACCGCCTCCTTG
AGCTGCTGCAGGAGGCCAGGTGCCAGGGAGGATTGTGCAGGGCCCTGAGCTGGGCCCGCACCCTGGAGCTGGGTG
CTGACTTCAGCCAGGTGCCCTCTGTGGACCATGTCCTGCACCAGCTAAAAGGTGTCCCCGAGGCCAACTTCTCCA
GCATGGTCCAGGAGGAGAACAGCACCTTCAACGCCCTTCCAGCCCTGGCTGCCATGCAGACATCCAGCGTGGTGC
AAGAGCTGAAGAAGGCAGTGGCCCAGCAGCCGGAAGGGGTGAGGACACTGGCTGAAGGGTTCCCGGGCTTGGAGG
CAGCTTCCCGCTGGGCCCAGGCACTGCAGGAGGTGGAGGAGAGCAGCCGCCCCTACCTGCAGGAGGTGCAGAGAT
ACGAGACCTACAGGTGGATCGTGGGCTGCGTGCTGTGCTCCGTGGTCCTATTCGTGGTGCTCTGCAACCTGCTGG
GCCTCAATCTGGGCATCTGGGGCCTGTCTGCCAGGGACGACCCCAGCCACCCAGAAGCCAAGGGCGAGGCTGGAG
CCCGCTTCCTCATGGCAGGTGTGGGCCTCAGCTTCCTCTTTGCTGCACCCCTCATCCTCCTGGTGTTCGCCACCT
TCCTGGTGGGTGGCAACGTGCAGACGCTGGTGTGCCGGAGCTGGGAGAACGGCGAGCTCTTTGAGTTTGCAGACA
CCCCCAGGGAACCTGCCCCCGTCCATGAACCTGTCGCAACTTCTTGGCCTGAGGAAGAACATCAGCATCCACCAAG
CCTATCAGCAGTGCAAGGAAGGGGCAGCGCTCTGGACAGTCCTGCAGCTCAACGACTCCTACGACCTGGAGGAGC
ACCTGGATATCAACCAGTATACCAACAAGCTACGGCAGGAGTTGCAGAGCCTGAAAGTAGACACACAGAGCCTGG
ACCTGCTGAGCTCAGCCGCCCGCCGGGACCTGGAGGCCCTGCAGAGCAGTGGGCTTCAGCGCATCCACTACCCCG
ACTTCCTCGTTCAGATCCAGAGGCCCGTGGTGAAGACCAGCATGGAGCAGCTGGCCCAGGAGCTGCAAGGACTGG
CCCAGGCCCAAGACAATTCTGTGCTGGGGCAGCGGCTGCAGGAGGAGGCCCAAGGACTCAGAAACCTTCACCAGG
AGAAGGTCGTCCCCCAGCAGAGCCTTGTGGCAAAGCTCAACCTCAGCGTCAGGGCCCTGGAGTCCTCTGCCCCGA
ATCTCCAGCTGGAGACCTCAGATGTCCTAGCCAATGTCACCTACCTGAAAGGAGAGCTGCCTGCCTGGGCAGCCA
GGATCCTGAGGAATGTGAGTGAGTGTTTCCTGGCCCGGGAGATGGGCTACTTCTCCCAGTACGTGGCCTGGGTGA
GAGAGGAGGTGACTCAGCGCATTGCCACCTGCCAGCCCCTCTCCGGAGCCCTGGACAACAGCCGTGTGATCCTGT
GTGACATGATGGCTGACCCCTGGAATGCCTTCTGGTTCTGCCTGGCATGGTGCACCTTCTTCCTGATCCCCACGA
TCATCTTTGCCGTCAAGACCTCCAAATACTTCCGTCCTATCCGGAAACGCCTCAGCTCCACCAGCTCTGAGGAGA
CTCAGCTCTTCCACATCCCCCGGGTTACCTCCCTGAAGCTGTAGGGCCTTGTGGGGTGAGGTGACCCTGAGGCTG
CCTGTCCTCCCCTTTGATTTAGCCTGGGCCACAGGACTTCGGTAGCTCTTGCCCCAGAGCCCAGGCTGGCATCCA
GGCCTGGACTGTCCCCAGTTCCGGCTTACCTGGCCCCACCTTGCCTGCTCCTTTCCACCCCTTTCTGCTCACGAC
CCCCATCATTCACGCTCAGAATCACATGGGACTTCTGTGCAGCTGCAGAGCCAGCAAGTCCCTACAGGTGTCACC
CGTTACCCCCATGCTGGTGGCATCCTCACAGGAAGAGCCTGTTCTCCACCTGCTGGAGCCTGGACCCTGGGGTGG
GACAGAGGCCTCGTCCAACCCCACTCCCCTTCCCGTGTGTCTTCCCCCTGCCAAGCCTCCCCCTGCCAAGCCTCC
CCCTGCCCCTCTCTGAGCCCCTCGCCCCCCACACCGTCCTCATCTGGCCTCCCCCCTGGCCCCCACTTCCCTCTT
ATGCCCTTCCTGGCCCTTTGCTTCCTCCCTTAGTCCCCTCTTCACCATATCTCCACTGCTACCTTGCTGGCCCCA
GAGACCACCCTGCCCAACCAAACCACTCAGGTAACGCCACTAATCAGGCAGGGGCCACCATGGCCTAGGTCTGGG
CTGGCTGCAGGCCCTGCCTCATGGCCTCTGAGCCCTCCACTGCCCCAGGGCCTTGGGCCCTCTGCAGATCTCATC
CAGGATTTATTGTTGTCCAGTGGGGTGAGGGAGGCCTGTCTGAAGGCCGAGCCTCCCTGCCTGCACCCAAGTTAG
AAATGGGGGTACCAGCACTTAGCTTCTCTCTGAGTGCTGGCTCCCAAGGAAGGGACCTGGGACCTGGGCCACAGT
GGGGGCTTGCCCTTACCTCTTCAGAAGGAAGCATCTTCCACAGCCCCCACCCAACTTTCTTAGGAGTGATCTGGT
GGCCAGAACAGGATTTTGCACGGCCCCTTTTATCCTGCGCATGTGGCCTAGGGTCATCCCCAGCCCATCCCTGTG
TCAGCCCTGAGTGCTGGACACTGCGTTCCAGAAATGAGGAAGAGGAGAGAAGAGATGGACAGACCTCAGATCC
ATTAAAGTGTTCTCACTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 148

MKHTLALLAPLLGLGLGLALSQLAAGATDCKFLGPAEHLTFTPAARARWLAPRVRAPGLL
DSLYGTVRRFLSVVQLNPFPSELVKALLNELASVKVNEVVRYEAGYVVCAVIAGLYLLLV
PTAGLCFCCCRCHRRCGGRVKTEHKALACERAALMVFLLLTTLLLLIGVVCAFVTNQRTH
EQMGPSIEAMPETLLSLWGLVSDVPQELQAVAQQFSLPQEQVSEELDGVGVSIGSAIHTQ
LRSSVYPLLAAVGSLGQVLQVSVHHLQTLNATVVELQAGQQDLEPAIREHRDRLLELLQE
ARCQGDCAGALSWARTLELGADFSQVPSVDHVLHQLKGVPEANFSSMVQEENSTFNALPA
LAAMQTSSVVQELKKAVAQQPEGVRTLAEGFPGLEAASRWAQALQEVEESSRPYLQEVQR
YETYRWIVGCVLCSVVLFVVLCNLLGLNLGIWGLSARDDPSHPEAKGEAGARTLMAGVGL
SFLFAAPLILLVFATFLVGGNVQTLVCRSWENGELFEFADTPGNLPPSMNLSQLLGLRKN
ISIHQAYQQCKEGAALWTVLQLNDSYDLEEHLDINQYTNKLRQELQSLKVDTQSLDLLSS
AARRDLEALQSSGLQRIHYPDFLVQIQRPVVKTSMEQLAQELQGLAQAQDNSVLGQRLQE
EAQGLRNLHQEKVVPQQSLVAKLNLSVRALESSAPNLQLETSDVLANVTYLKGELPAWAA
RILRNVSECFLAREMGYFSQYVAWVREEVTQRIATCQPLSGALDNSRVILCDMMADPWNA
FWFCLAWCTFFLIPSIIFAVKTSKYFRPIRKRLSSTSSEETQLFHIPRVTSLKL

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 105-125, 153-173, 428-449, 476-500, 778-797

N-glycosylation sites:

amino acids 270-273, 343-347, 352-356, 530-534, 540-546, 563-567, 684-688, 707-711, 725-729

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 811-815

Tyrosine kinase phosphorylation site.

amino acids 95-103

N-myristoylation sites.

amino acids 13-19, 15-21, 17-23, 26-32, 58-64, 124-130, 168-174, 228-234, 230-236, 320-326, 338-344, 393-399, 429-435, 446-452, 477-483, 500-506, 536-542, 644-650, 761-767

Phospholipase A2 histidine active site.

aminop acids 129-137

4Fe-4S ferredoxins, iron-sulfur binding region signature.

amino acids 126-138

Mitochondrial energy transfer proteins signature.

amino acids 80-89

FIGURE 149

CACAGCTCCCTTCCCAGGACGTGAAAATCTGCCTTCTCACC**ATG**AGGCTTCTAGTCCTTTCCA

GCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTG

CCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAA

CAAACCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCC

TTTGGGTGGTGCCTGGGGCACTCCCACAGGTG**TAG**CACTCCCAAAGCAAGACTCCAGACAGCG

GAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCAC

AGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAG

AACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACC

CCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCAT

GACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACC

CAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA

TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGC

TGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCA

FIGURE 150

MRLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPC
KLEPEPRLWVVPGALPQV

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 48-52

Amidation sites.

amino acids 23-27, 33-37

FIGURE 151

CACCGGAGGGCACGCAGCTGACGGAGCTGCGCTGCGTTCGCCTCGTTTGCCTCGCGCCCTCCA
CTGGAGCTGTTCGCGCCTCCCGGCTCCCACCGCAGCCCACCCGGCAGAGGAGTCGCTACCAGC
GCCCAGTGCGCTCTGTCAGTCCGCAAACTCCTTGCCGCCCGCCCCGGGCTGGGCACCAAATAC
CAGGCTACC**ATG**GTCTACAAGACTCTCTTCGCTCTTTGCATCTTAACTGCAGGATGGAGGGTA
CAGAGTCTGCCTACATCAGCTCCTTTGTCTGTTTCTCTTCCGACAAACATTGTACCACCGACC
ACCATCTGGACTAGCTCTCCACAAAACACTGATGCAGACACTGCCTCCCCATCCAACGGCACT
CACAACAACTCGGTGCTCCCAGTTACAGCATCAGCCCCAACATCTCTGCTTCCTAAGAACATT
TCCATAGAGTCCAGAGAAGAGGAGATCACCAGCCCAGGTTCGAATTGGGAAGGCACAAACACA
GACCCCTCACCTTCTGGGTTCTCGTCAACAAGCGGTGGAGTCCACTTAACAACCACGTTGGAG
GAACACAGCTCGGGCACTCCTGAAGCAGGCGTGGCAGCTACACTGTCGCAGTCCGCTGCTGAG
CCTCCCACACTCATCTCCCCTCAAGCTCCAGCCTCATCACCCTCATCCCTATCAACCTCACCA
CCTGAGGTCTTTTCTGCCTCCGTTACTACCAACCATAGCTCCACTGTGACCAGCACCCAACCC
ACTGGAGCTCCAACTGCACCAGAGTCCCCGACAGAGGAGTCCAGCTCTGACCACACACCCACT
TCACATGCCACAGCTGAGCCAGTGCCCCAGGAGAAAACACCCCCAACAACTGTGTCAGGCAAA
GTGATGTGTGAGCTCATAGACATGGAGACCACCACCACCTTTCCCAGGGTGATCATGCAGGAA
GTAGAACATGCATTAAGTTCAGGCAGCATCGCCGCCATTACCGTGACAGTCATTGCCGTGGTG
CTGCTGGTGTTTGGAGTTGCAGCCTACCTAAAAATCAGGCATTCCTCCTATGGAAGACTTTTG
GACGACCATGACTACGGGTCCTGGGGAAACTACAACAACCCTCTGTACGATGACTCC**TAA**CAA
TGGAATATGGCCTGGGATGAGGATTAACTGTTCTTTATTTATAAGTGCTTATCCAGTAGAATT
AATAAGTACCTGATGCGCATTGAACGACAATCTTAAGCCCTGTTTTGTTGGTATGGTTGTTTT
TGTTTTCCTCCCTCTCCTCTGGCTGCTACAACTTCCCCTTTCTGGTACAAGAAGAACCATTCT
TTAAAGGTGAGTGGAGGCTGATTTGCAGCTGAAGTGGGCCAGCCTTGCACCAGCCAGGCCAGA
CCACCATGGTGAAGGCTTCTTTCCCCACTGCAGGACCCACTTTGAGAAGGATCGAGGAGGAGG
ATTTGGGTTGTTTTGTTAGGGGTTACTTTCAGGGGAACATTTCATTTGTGTTATTTCTTAAAC
TTCTATTTAGGAAATTACATTAAGTATTAATGAGGGGAAAGGAAATGAGCTCTACGAGGATTT
CACCTTGCATGGGAGAGAGCAGGGTTTTCTCAGATTCCTTTTTAATCTCTATTTATCTGGTTG
TTTCTGACAGGATGCTGCCTGCTTGGCTCTACGAGCTGGAAAGCAGCTTCTTAGCTGCCTAAT
TAATGAAAGATGAAAATAGGAAGTGCCCTGGAGGGGGCCAGCAGGTCACGGGGCAGAATCTCT
CAGGTTGCTGTGGGATCTCAGTGTGCCCCTACCTGTTCTCCCCTCCAGGCCACCTGTCTCTGT
AAAGGATGTCTGCTCTGTTCAAAAGGCAGCTGGGATCCCAGCCCACAAGTGATCAGCAGAGTT
GCATTTCCAAAGAAAAAGGCTATGAGATGAGCTGAGTTATAGAGAGAAAGGGAGAGGCATGTA
CGGTGTGGGGAAGTGGAAGAGAAGCTGGCGGGGGAGAAGGAGGCTAACCTGCACTGAGTACTT
CATTAGGACAAGTGAGAATCAGCTATTGATAATGGCCAGAGATATCCACAGCTTGGAGGAGCC
CAGAGACTGTTTGCTTTATACCCACACAGCAACTGGTCCACTGCTTTACTGTCTGTTGGATAA
TGGCTGTAAAATGTTTAAAAC

FIGURE 152

MVYKTLFALCILTAGWRVQSLPTSAPLSVSLPTNIVPPTTIWTSSPQNTDADTASPSNGTHNN

SVLPVTASAPTSLLPKNISIESREEEITSPGSNWEGTNTDPSPSGFSSTSGGVHLTTTLEEHS

SGTPEAGVAATLSQSAAEPPTLISPQAPASSPSSLSTSPPEVFSASVTTNHSSTVTSTQPTGA

PTAPESPTEESSSDHTPTSHATAEPVPQEKTPPTTVSGKVMCELIDMETTTTFPRVIMQEVEH

ALSSGSIAAITVTVIAVVLLVFGVAAYLKIRHSSYGRLLDDHDYGSWGNYNNPLYDDS

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 258-278

N-glycosylation sites.

amino acids 58-61, 62-65, 80-83, 176-179

Casein kinase II phosphorylation sites.

amino acids 49-52, 85-88, 95-98, 100-103, 120-123, 121-124, 141-144, 164-167, 191-194, 195-198, 200-203

Tyrosine kinase phosphorylation site.

amino acids 289-296

N-myristoylation sites.

amino acids 59-64, 115-120, 128-133, 133-138, 257-262, 297-302

FIGURE 153

ACGTCACTGTCTTGAAGCAGCAGTAGCCTGGGAAGTGAGGCAGGAGGAATTGAGAGGCAGGAA

GGGNGCTGGAGACACAGCTGAGCCTGGAAATGAGAGTGGGCATCGCCGTGGTCATCATGACTC

CTCTGCGGCGTGGTCACC**ATG**TTGGTTCACTGTGTTGGGCTCTTATTGACGGGTCTCCTGCTA

GGCCTGACCTTGGGTGCCGGAGCCCTGCTGGCTTCTGAGCCTATCTACCAACCACCTTCAGCC

TGGGTGCCAGCTGGGGGGCTGGTGGGGCTGGCGCTGCTGGGAGCCCTGCTCACACTTCGGTGG

CCACGTCCATTCACAGTTCTGGGCACAACCCTGCTGGGTTCTGCAGTGCTTGTGGCCTGTGTT

GACTACTTCCTGGAGGGGCTGGCACTGGGGAGTTGGCTGGGCCAACGCCTGCAGACACTTCCA

GCCTTGCCTTCTCTCTGC**TGA**TATAGCTGGGTCTTACTGGGGATCTGGCCAGCCTTGGGGGCC

CTTGGAGCCCTGGCCCAGTGGAAGCTCGTGCCTGAGGAACATGGAGGCCACGCTAATGGGTCT

GTTCCTGGTTTCCCAGATGCATAAAGGAAGACATATCCCTCCCCTGGGCAGCAAGGCTACAAT

GGGAGGGAGGGAGAACATGGGAGCATGTGAATAAAATGGCATTAAATACTGAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 154

MLVHCVGLLLTGLLLGLTLGAGALLASEPIYQPPSAWVPAGGLVGLALLGALLTLRWPRPFTV
LGTTLLGSAVLVACVDYFLEGLALGSWLGQRLQTLPALPSLC

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 38-55, 60-78

N-myristoylation sites.

amino acids 7-13, 12-18, 16-22, 22-28, 41-47, 50-56, 84-90, 88-94

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 67-78

FIGURE 155

TGCAATTAAAGGAGTCGGGTCTCTAACTGTTGATCTGTTTTTTTCCCTTCTGAGCA**ATG**GAGC

TTACCATCTTTATCCTGAGACTGGCCATTTACATCCTGACATTTCCCTTGTACCTGCTGAACT

TTCTGGGCTTGTGGAGCTGGATATGCAAAAAATGGTTCCCCTACTTCTTGGTGAGGTTCACTG

TGATATACAACGAACAGATGGCAAGCAAGAAGCGGGAGCTCTTCAGTAACCTGCAGGAGTTTG

CGGGCCCCTCCGGGAAACTCTCCCTGCTGGAAGTGGGCTGTGGCACGGGGGCCAACTTCAAGT

TCTACCCACCTGGGTGCAGGGTGACCTGTATTGACCCCAACCCCAACTTTGAGAAGTTTTTGA

TCAAGAGCATTGCAGAGAACCGACACCTGCAGTTTGAGCGCTTTGTGGTAGCTGCCGGGGAGA

ACATGCACCAGGTGGCTGATGGCTCTGTGGATGTGGTGGTCTGCACCCTGGTGCTGTGCTCTG

TGAAGAACCAGGAGCGGATTCTCCGCGAGGTGTGCAGAGTGCTGAGACCGGGAGGGGCTTTCT

ATTTCATGGAGCATGTGGCAGCTGAGTGTTCGACTTGGAATTACTTCTGGCAACAAGTCCTGG

ATCCTGCCTGGCACCTTCTGTTTGATGGGTGCAACCTGACCAGAGAGAGCTGGAAGGCCCTGG

AGCGGGCCAGCTTCTCTAAGCTGAAGCTGCAGCACATCCAGGCCCCACTGTCCTGGGAGTTGG

TGCGCCCTCATATCTATGGATATGCTGTGAAA**TAG**TGTGAGCTGGCAGTTAAGAGCTGAATGG

CTCAAAGAATTTAAAGCTTCAGTTTTACATTTAAAATGCTAAGTGGGAGAAGAGAAACCTTTT

TTTTGGGGGGCGGTTTTTTTGGTTTGTTGTTGGTTTTTTTTTTTTTTTTGGCAGGAGAATCTC

TTGAACCCAGAAGGCGAAGGTTGCAGTGAACCGAGATCATGCCATTGTACTCTAGCCTGGGTG

ACAAGAGCAAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAGAAGTAGAGACAGGGAGAC

GGGGTCTCACTGTGTTGCCTAGGCCGGTCTTGAACTCCTGGGCTCAAGTGATTCTCCCACCTT

GACCTCCTAAATTGTTGGGATTACAGGTGTGAGACAGTGCACCTGGCCGAAATAGCTCAAGTT

TCTGAAAAACAAATCTGAATCTATTTGTTATTCTTAGCGTCACTGGTCTGGCTTTCAGAATTA

ACATACAAGGTTGCCACACCTAGTTCTGCCCAGCTTTATGTCTTTTATTCCAGTATTCCACCA

AAGTTTGTTTTCCTGCATTCCAGTTCTCAAGTCTTAAGATAAAGATTGTACTTGACAGTTTAG

TATATCCATAAAACTATTTGAGGTGGTTAAGGTTCTTGGGTTCATTTTCCTTAATACTTTGCT

GAATATTGTAGATTGTAGGCAATGAAAAGTCTACTAAATTAGGAAAACCTTGAATAATTAGG

TATCCTAGGTAAGAGCCCCTAAACATCAAGCAATCTGTGAGTCTGTAAAGAAATAAATATTTT

TTGGATTATTCTTATCTAATTCCACCCCTGTTGGAAGATGATTTCTTTGTTCTTTGCAACTAT

GGAAGCTGTGAAAATCATCACAAGTGCCTCTGAAAGCGAGTGTTAGGTTGGTTAGAGGGTTTA

ATATTTTCTGCAATGGTTTGTAGGAATTTTAATAAATGTAGTATATTTTCTGAGATGATTTTG

TAAAAGTACTATTTTAAATATCAAATCAACCAATAAATTCACATTTGTGTTAGGAACAAAA

FIGURE 156

MELTIFILRLAIYILTFPLYLLNFLGLWSWICKKWFPYFLVRFTVIYNEQMASKKRELFSNLQ

EFAGPSGKLSLLEVGCGTGANFKFYPPGCRVTCIDPNPNFEKFLIKSIAENRHLQFERFVVAA

GENMHQVADGSVDVVVCTLVLCSVKNQERILREVCRVLRPGGAFYFMEHVAAECSTWNYFWQQ

VLDPAWHLLFDGCNLTRESWKALERASFSKLKLQHIQAPLSWELVRPHIYGYAVK

Signal peptide:

amino acids 1-29

N-glycosylation site.

amino acids 203-207

N-myristoylation sites.

amino acids 78-84, 80-86, 91-97, 201-207

FIGURE 157

CCGCTGAGATGTACGAACTTCCGGTTCTCCGGGCAGCTGCCACTGCTGTAGCTTCTGCCACCT
GCCACGACCGGGCCTCTCCCTGGCGTTTGGTCACCTCTGCTTCATTCTCCACCGCGCCTATGG
TCCCTCTTGGAGCCAGCGTGGCGGGCCTGGCGGCTCCCGGGTGGTGAGAGAGCGGTCCGGGAA
CG**ATG**AAGGCCTCGCAGTGCTGCTGCTGTCTCAGCCACCTCTTGGCTTCCGTCCTCCTCCTGC
TGTTGCTGCCTGAACTAAGCGGGCCCCTGGCAGTCCTGCTGCAGGCAGCCGAGGCCGCGCCAG
GTCTTGGGCCTCCTGACCCTAGACCACGGACATTACCGCCGCTGCCACCGGGCCCTACCCCTG
CCAGCAGCCGGGCCGTGGTCTGGCTGAAGCTGCGGGGCCGCGGGGCTCCGAGGGAGGCAATG
GCAGCAACCCTGTGGCCGGGCTTGAGACGGACGATCACGGAGGGAAGGCCGGGGAAGGCTCGG
TGGGTGGCGGCCTTGCTGTGAGCCCCAACCCTGGCGACAAGCCCATGACCCAGCGGGCCCTGA
CCGTGTTGATGGTGGTGAGCGGCGCGGTGCTGGTGTACTTCGTGGTCAGGACGGTCAGGATGA
GAAGAAGAAACCGAAAGACTAGGAGATATGGAGTTTTGGACACTAACATAGAAAATATGGAAT
TGACACCTTTAGAACAGGATGATGAGGATGATGACAACACGTTGTTTGATGCCAATCATCCTC
GAAGA**TAA**GAATGTGCCTTTTGATGAAAGAACTTTATCTTTCTACAATGAAGAGTGGAATTTC
TATGTTTAAGGAATAAGAAGCCACTATATCAATGTTGGGGGGGTATTTAAGTTACATATATTT
TAACAACCTTTAATTTGCTGTTGCAATAAATACCGTATCCTTTTATTATATCTTTATATGTAT
AGAAGTACTCTATTAATGGGCTCAGAGATGTTGGGGATAAAGTATACTGTAATAATTTATCTG
TTTGAAAATTACTATAAAACGGTGTTTTCTGGTCGGTTTTGTTTCCTGCTTACCATATGATT
GTAAATTGTTTTATGTATTAATCAGTTAATGCTAATTATTTTGCTGATGTCATATGTTAAAG
AGCTATAAATTCCAACAACCAACTGGTGTGTAAAAATAATTTAAAATTTCCTTTACTGAAAGG
TATTTCCCATTTTTGTGGGGAAAAGAAGCCAAATTTATTACTTTGTGTTGGGGTTTTTAAAAT
ATTAAGAAATGTCTAAGTTATTGTTTGCAAAACAATAAATATGATTTTAAATTCTCTTAAAAA
AAAAA

FIGURE 158

MKASQCCCCLSHLLASVLLLLLLPELSGPLAVLLQAAEAAPGLGPPDPRPRTLPPLPPGPTPA
QQPGRGLAEAAGPRGSEGGNGSNPVAGLETDDHGGKAGEGSVGGGLAVSPNPGDKPMTQRALT
VLMVVSGAVLVYFVVRTVRMRRRNRKTRRYGVLDTNIENMELTPLEQDDEDDDNTLFDANHPRR

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 124-140

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 69-75, 78-84, 81-87, 97-103, 103-109, 106-112, 157-160

FIGURE 159

```
GCTGCAGGCGGCGACGGCTACACCATGGGCCGGCTGCTGCGGGCCGCCCGGCTGCCGCCGCTG
CTTTCGCCGCTGCTGCTTCTGCTGGTTGGGGGAGCGTTCCTGGGTGCCTGTGTGGCTGGGTCT
GATGAGCCTGGCCCAGAGGGCCTCACCTCCACCTCCCTGCTAGACCTCCTGCTGCCCACTGGC
TTGGAGCCACTGGACTCAGAGGAGCCTAGTGAGACCATGGGCCTGGGAGCTGGGCTGGGAGCC
TCTGGCTCAGGCTTCCCCAGCGAAGAGAATGAAGAGTCTCGGATTCTGCAGCCACCACAGTAC
TTCTGGGAAGAGGAGGAAGAGCTGAATGACTCAAGTCTGGACCTGGGACCCACTGCAGATTAT
GTTTTTCCTGACTTAACTGAGAAGGCAGGTTCCATTGAAGACACTAGCCAGGCTCAAGAGCTG
CCAAACCTCCCCTCTCCCTTGCCCAAGATGAATCTGGTTGAGCCTCCCTGGCATATGCCTCCC
AGAGAGGAGGAAGAAGAGGAAGAGGAAGAGGAGGAGAGGGAGAAGGAAGAGGTAGAGAAACAA
GAGGAGGAGGAAGAGGAGGAGCTGCTCCCTGTGAATGGATCCCAAGAAGAAGCCAAGCCTCAG
GTCCGTGACTTTTCTCTCACCAGCAGCAGCCAGACCCCAGGGGCCACCAAAAGCAGGCATGAA
GACTCCGGGGACCAGGCCTCATCAGGTGTGGAGGTGGAGAGCAGCATGGGGCCCAGCTTGCTG
CTGCCTTCAGTCACCCCAACTACAGTGACTCCGGGGGACCAGGACTCCACCAGCCAAGAGGCA
GAGGCCACAGTGCTGCCAGCTGCAGGGCTTGGGGTAGAGTTCGAGGCTCCTCAGGAAGCAAGC
GAGGAAGCCACTGCAGGAGCAGCTGGTTTGTCTGGCCAGCACGAGGAGGTGCCGGCCTTGCCT
TCATTCCCTCAAACCACAGCTCCCAGTGGGGCCGAGCACCCAGATGAAGATCCCCTTGGCTCT
AGAACCTCAGCCTCTTCCCCACTGGCCCCTGGAGACATGGAACTGACACCTTCCTCTGCTACC
TTGGGACAAGAAGATCTCAACCAGCAGCTCCTAGAAGGGCAGGCAGCTGAAGCTCAATCCAGG
ATACCCTGGGATTCTACGCAGGTGATCTGCAAGGACTGGAGCAATCTGGCTGGGAAAAACTAC
ATCATTCTGAACATGACAGAGAACATAGACTGTGAGGTGTTCCGGCAGCACCGGGGGCCACAG
CTCCTGGCCCTGGTGGAAGAGGTGCTGCCCCGCCATGGCAGTGGCCACCATGGGGCCTGGCAC
ATCTCTCTGAGCAAGCCCAGCGAGAAGGAGCAGCACCTTCTCATGACACTGGTGGGCGAGCAG
GGGGTGGTGCCCACTCAAGATGTCCTTTCCATGCTGGGTGACATCCGCAGGAGCCTGGAGGAG
ATTGGCATCCAGAACTATTCCACAACCAGCAGCTGCCAGGCGCGGGCCAGCCAGGTGCGCAGC
GACTACGGCACGCTCTTCGTGGTGCTGGTGGTCATTGGGGCCATCTGCATCATCATCATTGCG
CTTGGCCTGCTCTACAACTGCTGGCAGCGCCGGCTGCCCAAGCTCAAGCACGTGTCGCACGGC
GAGGAGCTGCGCTTCGTGGAGAACGGCTGCCACGACAACCCCACGCTGGACGTGGCCAGCGAC
AGCCAGTCGGAGATGCAGGAGAAGCACCCCAGCCTGAACGGCGGCGGGGCCCTCAACGGCCCG
GGGAGCTGGGGGGCGCTCATGGGGGGGCAAGCGGGACCCCGAGGACTCGGACGTGTTCGAGGAG
GACACGCACCTGTGAGCGCAGCCGAGGCGCAGGCCGAGTGGGCCGCCAGGACCAAGCGAGGTG
GACCCCGAAACGGACGGCCCGGAGCCCGCACCAGCCCCGCGCCTACCCGGGCCGCCCCCGCGG
CCTGGCCCTCGGCGCGGGCTCCTTCCCGCTTCCCCCGACTTCACACGGCGGCTTCGGACCAAC
TCCCTCACTCCCGCCCGAGGGGCAGGCCTCAAAGCCCGCCTTGGCCCCGCTTTCCCGCCCCTG
AACCCCGGCCCCGCGGGCGGCGGGCGGCGCTTCCTGCGCCCCGGGACTCAATTAAACCCGCCC
GGAGACCACGCCGGGCCCAGCAAAA
```

FIGURE 160

MGRLLRAARLPPLLSPLLLLLVGGAFLGACVAGSDEPGPEGLTSTSLLDLLLPTGLEPLDSEE
PSETMGLGAGLGASGSGFPSEENEESRILQPPQYFWEEEEELNDSSLDLGPTADYVFPDLTEK
AGSIEDTSQAQELPNLPSPLPKMNLVEPPWHMPPREEEEEEEEEEEEREKEEVEKQEEEEEEL
LPVNGSQEEAKPQVRDFSLTSSSQTPGATKSRHEDSGDQASSGVEVESSMGPSLLLPSVTPTT
VTPGDQDSTSQEAEATVLPAAGLGVEFEAPQEASEEATAGAAGLSGQHEEVPALPSFPQTTAP
SGAEHPDEDPLGSRTSASSPLAPGDMELTPSSATLGQEDLNQQLLEGQAAEAQSRIPWDSTQV
ICKDWSNLAGKNYIILNMTENIDCEVFRQHRGPQLLALVEEVLPRHGSGHHGAWHISLSKPSE
KEQHLLMTLVGEQGVVPTQDVLSMLGDIRRSLEEIGIQNYSTTSSCQARASQVRSDYGTLFVV
LVVIGAICIIIIALGLLYNCWQRRLPKLKHVSHGEELRFVENGCHDNPTLDVASDSQSEMQEK
HPSLNGGGALNGPGSWGALMGGKRDPEDSDVFEEDTHL

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 499-521

N-glycosylation sites.

amino acids 106-110, 193-197, 395-399, 480-484

Glycosaminoglycan attachment site.

amino acids 77-81

N-myristoylation sites.

amino acids 24-30, 28-34, 41-47, 69-75, 71-77, 73-79, 75-81, 216-222, 327-333, 455-461, 519-525, 574-580, 581-587, 584-590

Amidation site.

amino acids 588-592

FIGURE 161

CCAGGGCGGAGCGCAGCTGCGCCGGGCTTGGGCGCCTGGGGCCGCCGCTCCCCACCGTCGTTT

TCCCCACCGAGGCCGAGGCGTCCCGGAGTC**ATG**GCCGGCCTGAACTGCGGGGTCTCTATCGCA

CTGCTAGGGGTTCTGCTGCTGGGTGCGGCGCGCCTGCCGCGCGGGGCAGAAGCTTTTGAGATT

GCTCTGCCACGAGAAAGCAACATTACAGTTCTCATAAAGCTGGGGACCCCGACTCTGCTGGCA

AAACCCTGTTACATCGTCATTTCTAAAAGACATATAACCATGTTGTCCATCAAGTCTGGAGAA

AGAATAGTCTTTACCTTTAGCTGCCAGAGTCCTGAGAATCACTTTGTCATAGAGATCCAGAAA

AATATTGACTGTATGTCAGGCCCATGTCCTTTTGGGGAGGTTCAGCTTCAGCCCTCGACATCG

TTGTTGCCTACCCTCAACAGAACTTTCATCTGGGATGTCAAAGCTCATAAGAGCATCGGTTTA

GAGCTGCAGTTTTCCATCCCTCGCCTGAGGCAGATCGGTCCGGGTGAGAGCTGCCCAGACGGA

GTCACTCACTCCATCAGCGGCCGAATCGATGCCACCGTGGTCAGGATCGGAACCTTCTGCAGC

AATGGCACTGTGTCCCGGATCAAGATGCAAGAAGGAGTGAAAATGGCCTTACACCTCCCATGG

TTCCACCCCAGAAATGTCTCCGGCTTCAGCATTGCAAACCGCTCATCTATAAACGTCTGTGC

ATCATCGAGTCTGTGTTTGAGGGTGAAGGCTCAGCAACCCTGATGTCTGCCAACTACCCAGAA

GGCTTCCCTGAGGATGAGCTCATGACGTGGCAGTTTGTCGTTCCTGCACACCTGCGGGCCAGC

GTCTCCTTCCTCAACTTCAACCTCTCCAACTGTGAGAGGAAGGAGGAGCGGGTTGAATACTAC

ATCCCGGGCTCCACCACCAACCCCGAGGTGTTCAAGCTGGAGGACAAGCAGCCTGGGAACATG

GCGGGGAACTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAAGTCCAGGGATCCTC

CGGCTGCAGTTCCAAGTTTTGGTCCAACATCCACAAAATGAAAGCAGTGAG**TGA**GCCCCACTT

TCCTTTTTCTTCCTCCTCCAGCACCTTCGTTGTTTCCTGGGTAGTCTGCCTGGGTGAGGCTCC

CTTCCTGTTTCTCATCTGTGGCTTCTGAAACACTTAGACTCTGGACCCAGCAAGAGTTTCAGG

AAGTGGGTTGCTAGGCAGTTAGACAGGCTTGTTGGTGAACACCCGGTATGTAGTTCCATTTCA

GCACAATAAAAGAAATCTTGCATTCAAGATGCTAAATTGTTTTTAACGAAAA

FIGURE 162

MAGLNCGVSIALLGVLLLGAARLPRGAEAFEIALPRESNITVLIKLGTPTLLAKPCYIVISKR

HITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMSGPCPFGEVQLQPSTSLLPTLNRTFI

WDVKAHKSIGLELQFSIPRLRQIGPGESCPDGVTHSISGRIDATVVRIGTFCSNGTVSRIKMQ

EGVKMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFEGEGSATLMSANYPEGFPEDELMTW

QFVVPAHLRASVSFLNFNLSNCERKEERVEYYIPGSTTNPEVFKLEDKQPGNMAGNFNLSLQG

CDQDAQSPGILRLQFQVLVQHPQNESSE

Signal peptide:

amino acids 1-29

N-glycosylation sites.

amino acids 39-43, 122-126, 180-184, 205-209, 213-217, 270-274, 310-314, 339-343

Tyrosine kinase phosphorylation site.

amino acids 276-284

N-myristoylation sites.

amino acids 3-9, 7-13, 158-164, 175-181, 191-197, 303-309

FIGURE 163

CAACCACACACCTGGGGAATTGCTGGCCTGACTTCTGACCCCTGACTCCTCATACCCTTCCTC

CAGAGCATGACATTTGACCACCAACTGAAACCTGACCTCTGACCCCAGACCACTGGCCCTTCC

CCCGCCCTGTGGTGACTTCATAAAGGTTACTAGCTTCTCCCCTGGCCTTGAGACCCACACG**AT**

GGCCCTGCTGGCTCTGGCCAGTGCCGTCCCCTCTGCCCTGCTGGCCCTGGCTGTCTTCAGGGT

GCCCGCCTGGGCCTGTCTCCTCTGCTTCACAACCTACTCTGAGCGCCTCCGCATCTGCCAGAT

GTTTGTTGGGATGCGGAGCCCCAAGCTTGAAGAGTGTGAGGAGGCCTTCACGGCCGCCTTCCA

GGGCCTCTCTGACACCGAAATCAGTGAGGAGACCATCCACACTTCATCAGTGTCCTGGGGAAG

GTGCAGAGGGAGGGCAGGAGAGGCCCAGAGGGTCAGGCTGAGGGACAGACAGAGAGAAACAGT

CAGAGGAGAAAGGCTCAAAGACCATGAGAACAACAGAGACTTAGGGACAGAGAGACACAGACA

GGGGAAGACAGCAGGGCAAAGACTCAGAGAGGGGAGGATGGAGAGTCAGAGAGGGGAAGATGG

AGACTCAGAGAGAGGGGAGGATGGAGACTCAGAGAGAGAGGAAGATGGAGACTCAGAGGGAAA

GATGGAGACTCAGGAGTATGGAGAGTCAGAGAGGGGAGGATGGACACTCAGGGGAGGATGGAG

AGTCAGGAGGATGGAGACTCATAGAAAGGGGAGGATGGAGAGTCAGGAGAGGTTGGAGACTGG

AGAGGGAATAGAGACCCAGAAAGGGGAGGATGGAGACTCAGAGGGTGGAAGATGGAGACTCAA

AGAGGATGGAAACCCAGAGAGAGGAGGACAGAGA**TGA**GGCAGAGACTAGGGGAAGCAGGATAG

CGACTGGTCGGGGGCAGAGACTCAGGGAGGATAGAGACTCACAGAGAGGTGAGGATAGAGACT

TGGGAGGGACTCAGGAAGCATAGCGACTGTGGGGCAAAGAGTCAGAGAGGGGAGGATACAGAC

TTGGGAGGGCAGAGACTCAGAAACAGAATGTTCGCATTAGGGACATGGTGTTGCGGGGAGCTG

CCTCCCCCAGCCCCTGCTCCCTCCCTCACCGCCAGACTATGATGAGAGAAGCCACCTGCATGA

CACCTTCACCCAGATGACCCATGCCCTGCAGGAGCTGGCTGCTGCCCAGGGATCCTTTGAGGT

TGCCTTCCCTGATGCTGCAGAGAAAATGAAGAAGGTCATTACACAGCTTAAAGAAGCCCAGGC

TTGCATCCCTCCCTGCGGTCTCCAGGAGTTCGCCCGGCGTTTCCTCTGCAGCGGGTGCTACTC

TAGGGTCTGCGACCTCCCGCTGGACTGCCCAGTTCAGGATGTGACAGTGACTCGGGGCGACCA

GGCTATGTTTTCTTGCATCGTAAACTTCCAGCTGCCAAAGGAGGAGATCACCTATTCCTGGAA

GTTCGCAGGAGGAGGTCTCCGGACTCAGGACTTGTCCTATTTCCGAGATATGCCGCGGGCCGA

AGGATACCTGGCGCGGATCCGGCCGGCTCAGCTCACGCACCGCGGGACGTTCTCCTGCGTGAT

CAAGCAAGACCAGCGCCCCCTGGCCCGGCTCTACTTCTTTCTTAACGTCCTCGGGGCCCTCGC

ATCAGCGAGTGCGACAGTGTTGGCGTGGTGAGTTCTGGGGACTCCGGAGCCCCAGCATCTAGC

TCCCCGCTGTCTCAGATCCCACCGAGAAGTCTGGGTTCCCAGCAACCTCCAACCCAGGAGGAT

GTTCTTTCGATGGTACTGCAGTGGCAACTAACAAAGGTATCTTTCCTCCTTCCCTATCCTATT

TCCATCCTGAAAATAAAGAATATATTTCAACTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAA

FIGURE 164

MALLALASAVPSALLALAVFRVPAWACLLCFTTYSERLRICQMFVGMRSPKLEECEEAFTAAF
QGLSDTEISEETIHTSSVSWGRCRGRAGEAQRVRLRDRQRETVRGERLKDHENNRDLGTERHR
QGKTAGQRLREGRMESQRGEDGDSERGEDGDSEREEDGDSEGKMETQEYGESERGGWTLRGGW
RVRRMETHRKGRMESQERLETGEGIETQKGEDGDSEGGRWRLKEDGNPERGGQR

Signal peptide:

amino acids 1-26

N-myristoylation site.

amino acids 65-71

FIGURE 165

CAGAATCGCAGATTGCCAGCCCTTTTCCCGACCCCTACGGAAAGACGAGTCCAGGGGCCGTCC

TGGCGAGGTCAAAACATTTAGTCTGGTCTTTTCAGCGTGGACCCTGCCAGCAGCCAGGCC**ATG**

GAGCTCTCTGATGTCACCCTCATTGAGGGTGTGGGTAATGAGGTGATGGTGGTGGCAGGTGTG

GTGGTGCTGATTCTAGCCTTGGTCCTAGCTTGGCTCTCTACCTACGTAGCAGACAGCGGTAGC

AACCAGCTCCTGGGCGCTATTGTGTCAGCAGGCGACACATCCGTCCTCCACCTGGGGCATGTG

GACCACCTGGTGGCAGGCCAAGGCAACCCCGAGCCAACTGAACTCCCCATCCATCAGAGGGT

AATGATGAGAAGGCTGAAGAGGCGGGTGAAGGTCGGGGAGACTCCACTGGGGAGGCTGGAGCT

GGGGGTGGTGTTGAGCCCAGCCTTGAGCATCTCCTTGACATCCAAGGCCTGCCCAAAAGACAA

GCAGGTGCAGGCAGCAGCAGTCCAGAGGCCCCCCTGAGATCTGAGGATAGCACCTGCCTCCCT

CCCAGCCCTGGCCTCATCACTGTGCGGCTCAAATTCCTCAATGATACCGAGGAGCTGGCTGTG

GCTAGGCCAGAGGATACCGTGGGTGCCCTGAAGAGCAAATACTTCCCTGGACAAGAAAGCCAG

ATGAAACTGATCTACCAGGGCCGCCTGCTACAAGACCCAGCCCGCACACTGCGTTCTCTGAAC

ATTACCGACAACTGTGTGATTCACTGCCACCGCTCACCCCCAGGGTCAGCTGTTCCAGGCCCC

TCAGCCTCCTTGGCCCCCTCGGCCACTGAGCCACCCAGCCTTGGTGTCAATGTGGGCAGCCTC

ATGGTGCCTGTCTTTGTGGTGCTGTTGGGTGTGGTCTGGTACTTCCGAATCAATTACCGCCAA

TTCTTCACAGCACCTGCCACTGTCTCCCTGGTGGGAGTCACCGTCTTCTTCAGCTTCCTAGTA

TTTGGGATGTATGGACGA**TAA**GGACATAGGAAGAAAATGAAAGGCATGGTCTTTCTCCTTTAT

GGCCTCCCCACTTTTCCTGGCCAGAGCTGGGCCCAAGGGCCGGGGAGGGAGGGGTGGAAAGGA

TGTGATGGAAATCTCCTCCATAGGACACAGGAGGCAAGTATGCGGCCTCCCCTTCTCATCCAC

AGGAGTACAGATGTCCCTCCCGTGCGAGCACAACTCAGGTAGAAATGAGGATGTCATCTTCCT

TCACTTTTAGGGTCCTCTGAAGGAGTTCAAAGCTGCTGGCCAAGCTCAGTGGGGAGCCTGGGC

TCTGAGATTCCCTCCCACCTGTGGTTCTGACTCTTCCCAGTGTCCTGCATGTCTGCCCCCAGC

ACCCAGGGCTGCCTGCAAGGGCAGCTCAGCATGGCCCCAGCACAACTCCGTAGGGAGCCTGGA

GTATCCTTCCATTTCTCAGCCAAATACTCATCTTTTGAGACTGAAATCACACTGGCGGGAATG

AAGATTGTGCCAGCCTTCTCTTATGGGCACCTAGCCGCCTTCACCTTCTTCCTCTACCCCTTA

GCAGGAATAGGGTGTCCTCCCTTCTTTCAAAGCACTTTGCTTGCATTTTATTTTATTTTTTTA

AGAGTCCTTCATAGAGCTCAGTCAGGAAGGGGATGGGGCACCAAGCCAAGCCCCCAGCATTGG

GAGCGGCCAGGCCACAGCTGCTGCTCCCGTAGTCCTCAGGCTGTAAGCAAGAGACAGCACTGG

CCCTTGGCCAGCGTCCTACCCTGCCCAACTCCAAGGACTGGGTATGGATCGCTGGGCCCTAGG

CTCTTGCTTCTGGGGCTATTGGAGGGTCAGTGTCTGTGACTGAATAAAGTTCCATTTTGTGGA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 166

MELSDVTLIEGVGNEVMVVAGVVVLILALVLAWLSTYVADSGSNQLLGAIVSAGDTSVLHLGH

VDHLVAGQGNPEPTELPHPSEGNDEKAEEAGEGRGDSTGEAGAGGGVEPSLEHLLDIQGLPKR

QAGAGSSSPEAPLRSEDSTCLPPSPGLITVRLKFLNDTEELAVARPEDTVGALKSKYFPGQES

QMKLIYQGRLLQDPARTLRSLNITDNCVIHCHRSPPGSAVPGPSASLAPSATEPPSLGVNVGS

LMVPVFVVLLGVVWYFRINYRQFFTAPATVSLVGVTVFFSFLVFGMYGR

Signal peptide:

amino acids 1-36

Transmembrane domains:

amino acids 246-267, 275-301

N-glycosylation sites.

amino acids 162-166, 211-215

N-myristoylation sites.

amino acids 48-54, 105-111, 109-115, 129-135, 177-183, 247-253

Cell attachment sequence.

amino acids 97-100

FIGURE 167

GGCGGCTGTGTGTCGCCGGAGCCGAAGCGCGCAGGCCCGTCCCGGTGGCCGGGGAGCGGGCGGGTGGGGGCGCC**A**
TGTGGTTCATGTACCTGCTGAGCTGGCTGTCGCTCTTCATCCAGGTGGCCTTCATCACGCTGGCTGTCGCGGCTG
GACTCTATTACCTGGCAGAACTGATAGAAGAATACACAGTGGCCACCAGCAGGATCATAAAATACATGATCTGGT
TCTCCACCGCTGTACTGATTGGCCTCTACGTCTTTGAGCGCTTCCCCACCAGCATGATTGGAGTGGGCCTATTCA
CCAACCTCGTCTACTTTGGCCTCCTCCAGACCTTCCCCTTCATCATGCTGACCTCGCCTAACTTCATCCTGTCGT
GTGGACTAGTGGTGGTGAATCATTACCTAGCATTTCAGTTTTTTGCAGAAGAATATTATCCCTTCTCAGAGGTCC
TGGCCTATTTCACTTTCTGCCTGTGGATAATTCCGTTTGCGTTTTTTGTGTCACTTTCGGCCGGGGAGAACGTCC
TGCCCTCTACCATGCAGCCAGGAGATGATGTCGTCTCCAATTATTTCACCAAAGGCAAGCGGGGCAAACGCTTAG
GGATCCTGGTTGTCTTCTCCTTCATCAAAGAGGCCATTCTACCCAGTCGTCAGAAGATATAC**TGA**CCCCCATGCA
GGCAGGATGTGGGGGGCAAGATCAGGAGAGTCAGGCCCCTGGGCCTCTATGCCAGGTGGGGACCAGAAGTCGGGA
AGGCACCTACCACCTGCCCTGGCTTTCTTCCCCTCAACTCTGGAGCCCCATCCCCACCCTCCTTGGGGGGCTCAG
CTTGGCTCAGATCTGATGCTTCAAGAGGCTGTAACCTCAGAGGGCACCAAGGAGGGTGGCAGAGCCTGCTTAGCC
AGGAGGCCGAGGTCCCTCAGTCCTCCCCTGTCCCTTCCAAGGTGGGTCAGGAGGTTCTGGCCCCGCTGGGGCAGG
CAGGGCAGGGTCTGTGAAGCTTAAGAGCAGATGGTGACAAGTTCTCTGGGCAGGTGGCCATGGGGAGGGGCCATG
GCTTGGCATGTCCAACAGAAATAGTTTTTGCTGTTGAACGGTGATTTCTGTCCAAGTGCAGATTTCCGTTTGAAT
AAAGCTTCGCTTCTAGGTGGCACTGTTTGCCTTAATACCCTGACAGTTCATCTTCCTTTCTTCCTGCTAACCTTC
TGCTCTGGACTGGACTCACTTTTCTGCTCCAGGGACTCCTTTTCTGGGTTTGGGTCTTGCCCTTCCCAAGGGACT
GTTCTTGTGGCCCTTAATGGGAAGGGGGCAGGGGTGAGGAGCTGAGCCTGCTCAAGGAGTGGGAAGTGGGGCTAT
AGGCAGCCTCTCTGATGCACTCTCTTCCATCTCTTTCCCCAAGGCTCCGTGACTGTCAAACTGGGAGTAGGAGAG
GGGACAATTTAGGACTGGGCTAGATTTTCAGAAGAACATCTACAATATCCTATTTATAAATCTTCCTCTGGGAAA
AGGAGTGGTTTCTGGCTGAATACTATCTTAGGCTCAAGGAGAAACAAAATAAAAATTAGCTTCCAGGCAGCCTGT
TTTTAAAGAAATGGGACTAATGGGAGAAGCTGTTTGTCACTCTAAGAGCATCCAAGCCCTGGCCCGTCTGTGCAC
TCTTGGCTCCTGGGGAGATATATCTGCCTTCTAAGAAGGCAGGCCAGGTCTTGGGCACAGACCTGCATTTGTTGA
CCTTGCACTCCAACTATAGTGCCTTGCAAGTGCTCAACAGTACATATTGGAATGAAGTCCCTATGAGAGCCATTT
CTGGCCATGTTCTATACCTCAAAGTGAGGCTGGCAGGTACAGAGATGAACTGTACACATGTGATACATTTAAGCC
ACTGGAAAAACCCCTGTGCTTGAAAATATTTCCTCTATATCATGCCTGGAGTTCCATCATAGCCCTTCATTTCCT
TGGCTTTAGCATTTACCTTCTCTTAAGAATACCAGCTTTCCCCTTTCCCTGAGAGGAAGAGCACATGTTGGTCTC
CTCTTAGTGTGAACGAGATTGCCAGGCCCTTTTCTCCTATGCACACCAGGATAGACAAGGCAGGGGATACTGGCA
GCCTGCATCATCCTCCCATTGGGCTGACAGCTGGCCCTACTTTCCTCCCTCTGCTGCTTGGTCCCTCACCTTGAT
GATGTGGCTTCGCCCCCTCCACTCTACTGCCAGTGTTCTCCCAGGGGTTGCTAAATCCAGCAGACCCCTTTCCTG
TCTTACTAGATCTGGGCAGCATTTGACATGGCTGATCACCCCTTGCTTCTTGGATGGCACTTCCCTGGCACCTCT
GTGGCTAGTTGTCCTACCTCCCTGGCTGTTCCTTTCAGGCTTCCGTGCAGGCTTCTCCACTTGCCCATGCACAGT
AGGGTCTTTCAGGGTTCTGCTGTGGGCTCCCTAGGGAAGCCCATCCATCTGGATGGTTTCAAGGATGGTGAGGAA
TTTAGAGTTGACCTCCAGCCCCAACATCCTTCCTGATCACCTGAACCACAGTTTTGCTGCCCTCTAGGTGCACAG
ACAATTCAGGTCCATGGCCCAGATGGTACTTGCTGTCTTCTGCAAACCTGCCCCTTCTGGGTACTTCCCTTGACC
CCGAGATCACTCAGGAGCCAGACAGGAAACTTATTCTATTCCTGTTTTCTCTTTCTGCCCACCACATCCAATCTC
TCAAAACGGTCAGGTCTACCTTAACATCTCTTGATTTGAGCCACTCCCACTGTCATCAGCTTTCACCTGGATTAT
CGTGACAGCCTCCTACTGCTTCTCTATCATGTGGCCAGAGCTATCTTCCTAAAATGCATTGCATAGTTGATCAAG
TCACTCTCTGGCCTAAAACCTTCCTTGGCTCCCTGCTGCCCTCAGGATAAAGTCTGGACCCCTCAGCATGGCTTG
TGAGACTCATGGTGTCCTTGTCCCTGCTCACCTCTCTGGTCTCATCACTTGCCTTCTTGCATTCTGGGTCCCAGC
CTCCTGTATCCAGAGATGCAGTGGCTCTCCATTGCCACTCTGATTCCTCCTTTCTTTTGGTCACAGAGAAAGGGT
ACTTTCTCTGTCAAATCTCAACTTAGACTTGACTTCCTCCAAGGAGCTTTGGCTATACTCTCTCCTCCCGACCCC
CACCCTGGCATACTACACAGATCACTCTGGGCTCACTTGCCTGCCTAATGGTCATCTCCCCAGTAGACTGTAAGC
TCCTTGAGGGCAAGGATTGTGTTGGAATTTTTGTATTAACAGTGCCTGGCTTGGTGCCTGGCACCTAGAAAGCAC
TCAATAAATGTTTGTTTAATGAA

FIGURE 168

MWFMYLLSWLSLFIQVAFITLAVAAGLYYLAELIEEYTVATSRIIKYMIWFSTAVLIGLYVFE
RFPTSMIGVGLFTNLVYFGLLQTFPFIMLTSPNFILSCGLVVVNHYLAFQFFAEEYYPFSEVL
AYFTFCLWIIPFAFFVSLSAGENVLPSTMQPGDDVVSNYFTKGKRGKRLGILVVFSFIKEAIL
PSRQKIY

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 126-146

Casein kinase II phosphorylation site.

amino acids 145-148

N-myristoylation sites.

amino acids 73-78, 82-87

Amidation sites.

amino acids 168-171, 171-174

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 91-101

FIGURE 169

CAAAGCCCTACCCTCACCATTCACCAGGTCCTGTGGGAAGAGCAGCGTGGAGGTGGGCTGAGG

TTAGAAGGTGCAGAGCGTGGAAGAAGATTGTGAGCTGAGTATTGGACATCTGTTCTTGAATAG

TCCCTGGGCCTGCCATAGGAAAGGAAGTTCTCCAGGGTTACAGTTCTTATCCGCGTGAATACA

C**ATG**GCTCTGTTACGAAAAATTAATCAGGTGCTGCTGTTCCTTCTGATCGTGACCCTCTGTGT

GATTCTGTATAAGAAAGTTCATAAGGGGACTGTGCCCAAGAATGACGCAGATGATGAATCCGA

GACTCCTGAAGAACTGGAAGAAGAGATTCCTGTGGTGATTTGTGCTGCAGCAGGGAGGATGGG

TGCCACTATGGCTGCCATCAATAGCATCTACAGCAACACTGACGCCAACATCTTGTTCTATGT

AGTGGGACTCCGGAATACTCTGACTCGAATACGAAAATGGATTGAACATTCCAAACTGAGAGA

AATAAACTTTAAAATCGTGGAATTCAACCCGATGGTCCTCAAAGGGAAGATCAGACCAGACTC

ATCGAGGCCTGAATTGCTCCAGCCTCTGAACTTTGTTCGATTTTATCTCCCTCTACTTATCCA

CCAACACGAGAAAGTCATCTATTTGGACGATGATGTAATTGTACAAGGTGATATCCAAGAACT

GTATGACACCACCTTGGCCCTGGGCCACGCGGCGGCTTTCTCAGATGACTGCGATTTGCCCTC

TGCTCAGGACATAAACAGACTCGTGGGACTTCAGAACACATATATGGGCTATCTGGACTACCG

GAAGAAGGCCATCAAGGACCTTGGCATCAGCCCCAGCACCTGCTCTTTCAATCCTGGTGTGAT

TGTTGCCAACATGACAGAATGGAAGCACCAGCGCATCACCAAGCAATTGGAGAAATGGATGCA

AAAGAATGTGGAGGAAAACCTCTATAGCAGCTCCCTGGGAGGAGGGGTGGCCACCTCCCCAAT

GCTGATTGTGTTTCATGGGAAATATTCCACAATTAACCCCCTGTGGCACATAAGGCACCTGGG

CTGGAATCCAGATGCCAGATATTCGGAGCATTTTCTGCAGGAAGCTAAATTACTCCACTGGAA

TGGAAGACATAAACCTTGGGACTTCCCTAGTGTTCACAACGACTTATGGGAAAGCTGGTTTGT

TCCTGACCCTGCAGGGATATTTAAACTCAATCACCATAGC**TGA**TATAACTCTACCCTTAAAAT

ATTCCCTGTATAGAAATGTGGAATTGTCCCTTTGTAGCCAACTATAACATTGTTCTTTATGAA

TATTACCTTTGATACATATGATCCACAATATAAAAACCAAAAACTACTGTGTGCAAATTATAC

CTTGGACCATATAGGCATTGATTAACTTCTTTAAGTACATGTGATAACTATGGAAATCAAGAT

TATGTGACTGAAAAACATAAAGGAAGAGACCCATCTAGATAACAGCAATCAACCTGCTTAATT

CTGAATGACAATTATATCCACAAATTTTTAAAACTTCTACATGTATTTTTCACATGAAGATCT

CCTTAACAGGTTGCCAACCTTTTCTTTTATAAAACTATTACATTTAAAATATGGACGTCTGAA

AAATAAAATATTCATCATTTTTAAAA

FIGURE 170

MALLRKINQVLLFLLIVTLCVILYKKVHKGTVPKNDADDESETPEELEEEIPVVICAAAGRMG

ATMAAINSIYSNTDANILFYVVGLRNTLTRIRKWIEHSKLREINFKIVEFNPMVLKGKIRPDS

SRPELLQPLNFVRFYLPLLIHQHEKVIYLDDDVIVQGDIQELYDTTLALGHAAAFSDDCDLPS

AQDINRLVGLQNTYMGYLDYRKKAIKDLGISPSTCSFNPGVIVANMTEWKHQRITKQLEKWMQ

KNVEENLYSSSLGGGVATSPMLIVFHGKYSTINPLWHIRHLGWNPDARYSEHFLQEAKLLHWN

GRHKPWDFPSVHNDLWESWFVPDPAGIFKLNHHS

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 234-238

Tyrosine kinase phosphorylation site.

amino acids 253-261

N-myristoylation sites.

amino acids 63-69, 86-92, 198-204, 218-224, 229-235, 265-271, 266-272

FIGURE 171

GCCAGAGGCTGCAGCTGGAGCCCAGAGCCCAAG**ATG**GAGCCCCAGCTGGGGCCTGAGGCTGCC

GCCCTCCGCCCTGGCTGGCTGGCCCTGCTGCTGTGGGTCTCAGCCCTGAGCTGTTCTTTCTCC

TTGCCAGCTTCTTCCCTTTCTTCTCTGGTGCCCCAAGTCAGAACCAGCTACAATTTTGGAAGG

ACTTTCCTCGGTCTTGATAAATGCAATGCCTGCATCGGGACATCTATTTGCAAGAAGTTCTTT

AAAGAAGAAATAAGATCTGACAACTGGCTGGCTTCCCACCTTGGACTGCCTCCCGATTCCTTG

CTTTCTTATCCTGCAAATTACTCAGATGATTCCAAAATCTGGCGCCCTGTGGAGATCTTTAGA

CTGGTCAGCAAATATCAAAACGAGATCTCAGACAGGAGAATCTGTGCCTCTGCATCAGCCCCA

AAGACCTGCAGCATTGAGCGTGTCCTGCGGAAAACAGAGAGGTTCCAGAAATGGCTGCAGGCC

AAGCGCCTCACGCCGGACCTGGTGCAGGACTGTCACCAGGGCCAGAGAGAACTAAAGTTCCTG

TGTATGCTGAGA**TAA**CACCAGTGAAAAGCCTGGCATGGAGCCCAGCACTGAGAACTTCCAGA

AAGTGTTAGCCTTCTCCCAACTGTGTTATACCAACCACATTTTCAAATAGTAATCATTAAAGA

GGCTTCTGCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 172

MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLVPQVRTSYNFGRTFLGLDKCNAC
IGTSICKKFFKEEIRSDNWLASHLGLPPDSLLSYPANYSDDSKIWRPVEIFRLVSKYQNEISD
RRICASASAPKTCSIERVLRKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCMLR

Signal peptide:
amino acids 1-28

N-glycosylation site.
amino acids 100-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 158-161

N-myristoylation sites.
amino acids 56-61, 65-70

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 18-28

Prenyl group binding site (CAAX box).
amino acids 179-182

Leucine zipper pattern.
amino acids 5-26

FIGURE 173

GCTGGACTGCTCGCTGGCCGGCAGCGCACCGTTTTGAAGGTCCTAGCCCACCTGGGCTGGCTC

ACGCGCACGACTAGCCGCTCCCATACAGCACGCCCGGACTCTGTCGTCGCTTAAGGCCACTCC

TATTCTACGGCTGACCCCTGGTGGTCACGTGGATCTGTTCGCCACGCAAGTCTGGGTCCTTCG

GCGATTGACCGGGGTCCTTGCTGTTCGGGAGCCTCTCCTAAGCTGCCTGTTCGCGCGAGAGTT

TGGAGGGGCGGGTTTGGGGTCGGTGTCTGATTGGGGCTCGCACCGCAGCACGCTGGAGTCCCG

CTTAGGTACCAGTTAGCGTCAGGGGAGCTGGGTCAGGCGGTCGCCGGGACACCCGTGTGTGG

CAGGCGGCGAAGCGCTCTGGAGAATCCCGGACAGCCCTGCTCCCTGCAGCCAGGTGTAGTTTC

GGGAGCCACTGGGGCCAAAGTGAGAGTCCAGCGGTCTTCCAGCGCTTGGGCCACGGCGGCGGC

CCTGGGAGCAGAGGTGGAGCGACCCCATTACGCTAAAG**ATG**AAAGGCTGGGGTTGGCTGGCCC

TGCTTCTGGGGGCCCTGCTGGGAACCGCCTGGGCTCGGAGGAGCCAGGATCTCCACTGTGGAG

CATGCAGGGCTCTGGTGGATGAACTAGAATGGGAAATTGCCCAGGTGGACCCCAAGAAGACCA

TTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAGTGGTGGAGGTGCCTTATG

CCCGCTCAGAGGCCCACCTCACAGAGCTGCTGGAGGAGATATGTGACCGGATGAAGGAGTATG

GGGAACAGATTGATCCTTCCACCCATCGCAAGAACTACGTACGTGTAGTGGGCCGGAATGGAG

AATCCAGTGAACTGGACCTACAAGGCATCCGAATCGACTCAGATATTAGCGGCACCCTCAAGT

TTGCGTGTGAGAGCATTGTGGAGGAATACGAGGATGAACTCATTGAATTCTTTTCCCGAGAGG

CTGACAATGTTAAAGACAAACTTTGCAGTAAGCGAACAGATCTTTGTGACCATGCCCTGCACA

TATCGCATGATGAGCTA**TGA**ACCACTGGAGCAGCCCACACTGGCTTGATGGATCACCCCCAGG

AGGGGAAAATGGTGGCAATGCCTTTTATATATTATGTTTTTACTGAAATTAACTGAAAAAATA

TGAAACCAAAAGT

FIGURE 174

MKGWGWLALLLGALLGTAWARRSQDLHCGACRALVDELEWEIAQVDPKKTIQMGSFRINPDGS

QSVVEVPYARSEAHLTELLEEICDRMKEYGEQIDPSTHRKNYVRVVGRNGESSELDLQGIRID

SDISGTLKFACESIVEEYEDELIEFFSREADNVKDKLCSKRTDLCDHALHISHDEL

Signal peptide:

amino acids 1-20

N-myristoylation sites.

amino acids 12-18, 16-22, 29-35

Endoplasmic reticulum targeting sequence.

amino acids 179-184

FIGURE 175

CGCAGCGCGGCAGTCCTG**ATG**GCCCGGCATGGGTTACCGCTGCTGCCCCTGCTGTCGCTCCTG

GTCGGCGCGTGGCTCAAGCTAGGAAATGGACAGGCTACTAGCATGGTCCAACTGCAGGGTGGG

AGATTCCTGATGGGAACAAATTCTCCAGACAGCAGAGATGGTGAAGGGCCTGTGCGGGAGGCG

ACAGTGAAACCCTTTGCCATCGACATATTTCCTGTCACCAACAAAGATTTCAGGGATTTTGTC

AGGGAGAAAAGTATCGGACAGAAGCTGAGATGTTTGGATGGAGCTTTGTCTTTGAGGACTTT

GTCTCTGATGAGCTGAGAAACAAAGCCACCCAGCCAATGAAGTCTGTACTCTGGTGGCTTCCA

GTGGAAAAGGCATTTTGGAGGCAGCCTGCAGGTCCTGGCTCTGGCATCCGAGAGAGACTGGAG

CACCCAGTGTTACACGTGAGCTGGAATGACGCCCGTGCCTACTGTGCTTGGCGGGGAAAACGA

CTGCCCACGGAGGAAGAGTGGGAGTTTGCCGCCCGAGGGGGCTTGAAGGGTCAAGTTTACCCA

TGGGGGAACTGGTTCCAGCCAAACCGCACCAACCTGTGGCAGGGAAAGTTCCCCAAGGGAGAC

AAAGCTGAGGATGGCTTCCATGGAGTCTCCCCAGTGAATGCTTTCCCCGCCCAGAACAACTAC

GGGCTCTATGACCTCCTGGGGAACGTGTGGGAGTGGACAGCATCACCGTACCAGGCTGCTGAG

CAGGACATGCGCGTCCTCCGGGGGGCATCCTGGATCGACACAGCTGATGGCTCTGCCAATCAC

CGGGCCCGGGTCACCACCAGGATGGGCAACACTCCAGATTCAGCCTCAGACAACCTCGGTTTC

CGCTGTGCTGCAGACGCAGGCCGGCCGCCAGGGGAGCTG**TAA**GCAGCCGGGTGGTGACAAGGA

GAAAAGCCTTCTAGGGTCACTGTCATTCCCTGGCCATGTTGCAAACAGCGCAATTCCAAGCTC

GAGAGCTTCAGCCTCAGGAAAGAACTTCCCCTTCCCTGTCTCCCATCCCTCTGTGGCAGGCGC

CTCTCACCAGGGCAGGAGAGGACTCAGCCTCCTGTGTTTTGGAGAAGGGGCCCAATGTGTGTT

GACGATGGCTGGGGGCCAGGTGTTTCTGTTAGAGGCCAAGTATTATTGACACAGGATTGCAAA

CACACAAACAGTTGGAACAGAGCACTCTGAAAGGCCATTTTTTAAGCATTTTAAAATCTATTC

TCTCCCCCTTTCTCCCTGGATGATTCAGGAAGCTGACATTGTTTCCTCAAGGCAGAATTTTCC

TGGTTCTGTTTTCTCAGCCAGTTGCTGTGGAAGGAGAATGCTTTCTTTGTGGCCTCATCTGTG

GTTTCGTGTCCCTCTGAAGGAAACTAGTTTCCACTGTGTAACAGGCAGACATGTAACTATTTA

AAGCACAGTTCAGTCCTAAAAGGGTCTGGGAGAACCAGATGATGTACTAGGTGAAGCATTGCA

TTGTGGGAATCACAAAGCAAATAGTACTCCAGAAAGACAAATATCAGAAGCTTCCTATTCTTT

TTTTTTTTTTTTTTTTTTTTTGAGACAGGGTCTTTCTCTGTTGCCCAGGCTAGAGTGCACTG

GTGATCACGGCTCACTCTAGCCTTGAATTCCTGGGCCCAAGCAATTCTCCCACCTCAGCCTCC

TGAGTAGCTGGGACTACAAGTGTGCACCACCATGCCTGGCTAATTTTTTGAATTTTTGTAGTG

ATGGGATCTCGCTCTGTTGCCCAGGGTGGTCTCGAACTCCTGGCCTCAAGCGATCCTCCCACC

TCGACCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCTCGCCTGGGCCCCCTTCTCCATA

TGCCTCCAAAAACATGTCCCTGGAGAGTAGCCTGCTCCCACACTGTCACTGGATGTCATGGGG

CCAATAAAATCTCCTGCAATTGTGTATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAA

FIGURE 176

MARHGLPLLPLLSLLVGAWLKLGNGQATSMVQLQGGRFLMGTNSPDSRDGEGPVREATVKPFA
IDIFPVTNKDFRDFVREKKYRTEAEMFGWSFVFEDFVSDELRNKATQPMKSVLWWLPVEKAFW
RQPAGPGSGIRERLEHPVLHVSWNDARAYCAWRGKRLPTEEEWEFAARGGLKGQVYPWGNWFQ
PNRTNLWQGKFPKGDKAEDGFHGVSPVNAFPAQNNYGLYDLLGNVWEWTASPYQAAEQDMRVL
RGASWIDTADGSANHRARVTTRMGNTPDSASDNLGFRCAADAGRPPGEL

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 191-195

N-myristoylation sites.

amino acids 23-29, 25-31, 175-181

Amidation site.

amino acids 159-163

FIGURE 177

GCCTTCTCGCGCCTGACCATGCACCCCTGCATCTTCCTGCTGGGCCACAGGCGAGCGCTTTAT

TTCTGGAGCTGAGGGCTAAAACTTTTTTGACTTTTCTTCTCCTCAACATCTGAATC**ATG**CCAT

GTGCCCAGAGGAGCTGGCTTGCAAACCTTTCCGTGGTGGCTCAGCTCCTTAACTTTGGGGCGC

TTTGCTATGGGAGACAGCCTCAGCCAGGCCCGGTTCGCTTCCCGGACAGGAGGCAAGAGCATT

TTATCAAGGGCCTGCCAGAATACCACGTGGTGGGTCCAGTCCGAGTAGATGCCAGTGGGCATT

TTTTGTCATATGGCTTGCACTATCCCATCACGAGCAGCAGGAGGAAGAGAGATTTGGATGGCT

CAGAGGACTGGGTGTACTACAGAATTTCTCACGAGGAGAAGGACCTGTTTTTTAACTTGACGG

TCAATCAAGGATTTCTTTCCAATAGCTACATCATGGAGAAGAGATATGGGAACCTCTCCCATG

TTAAGATGATGGCTTCCTCTGCCCCCCTCTGCCATCTCAGTGGCACGGTTCTACAGCAGGGCA

CCAGAGTTGGGACGGCAGCCCTCAGTGCCTGCCATGGACTGACTGGATTTTTCCAACTACCAC

ATGGAGACTTTTTCATTGAACCCGTGAAGAAGCATCCACTGGTTGAGGGAGGGTACCACCCGC

ACATCGTTTACAGGAGGCAGAAAGTTCCAGAAACCAAGGAGCCAACCTGTGGATTAAAGGGTA

TTGTGACTCACATGTCCTCCTGGGTTGAAGAATCTGTTTTGTTCTTTTGG**TAG**TTTTATTAAA

ACATGACCTATTCTTACTCAAGTCTCTTATCTCCTCTGTATTCTTTTTTTTTTAATATCTTCA

TGACATTCAAATCTCTTCTGTATTCTCTTGCCAGAAAGTGTACATTCTTTTTGCTTGTATAAA

CCCTTTCACTTGTC

FIGURE 178

MPCAQRSWLANLSVVAQLLNFGALCYGRQPQPGPVRFPDRRQEHFIKGLPEYHVVGPVRVDAS
GHFLSYGLHYPITSSRRKRDLDGSEDWVYYRISHEEKDLFFNLTVNQGFLSNSYIMEKRYGNL
SHVKMMASSAPLCHLSGTVLQQGTRVGTAALSACHGLTGFFQLPHGDFFIEPVKKHPLVEGGY
HPHIVYRRQKVPETKEPTCGLKGIVTHMSSWVEESVLFFW

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 11-15, 105-109, 125-129

N-myristoylation site.

amino acids 149-155

FIGURE 179

CAGATTTAAAAAGAAAACCTTTACTGAATCAGCTGAGTGTTAATAATACGAATTTCCTTTTCT
TGCCAATTCTGATCTGAACAGAAAATCCAAGAACAGGGAT**ATG**TGTGGATTACAGTTTTCTCT
GCCTTGCCTACGACTGTTTCTGGTTGTTACCTGTTATCTTTTATTATTACTCCACAAAGAAAT
ACTTGGATGTTCGTCTGTTTGTCAGCTCTGCACTGGGAGACAAATTAACTGCCGTAACTTAGG
CCTTTCGAGTATTCCTAAGAATTTTCCTGAAAGTACAGTTTTTCTGTATCTGACTGGGAATAA
TATATCTTATATAAATGAAAGTGAATTAACAGGACTTCATTCTCTTGTAGCATTGTATTTGGA
TAATTCTAACATTCTGTATGTATATCCAAAAGCCTTTGTTCAATTGAGGCATCTATATTTTCT
ATTTCTAAATAATAATTTCATCAAACGCTTAGATCCTGGAATATTTAAGGGACTTTTAAATCT
TCGTAATTTATATTTACAGTATAATCAGGTATCTTTTGTTCCGAGAGGAGTATTTAATGATCT
AGTTTCAGTTCAGTACTTAAATCTACAAAGGAATCGCCTCACTGTCCTTGGGAGTGGTACCTT
TGTTGGTATGGTTGCTCTTCGGATACTTGATTTATCAAACAATAACATTTTGAGGATATCAGA
ATCAGGCTTTCAACATCTTGAAAACCTTGCTTGTTTGTATTTAGGAAGTAATAATTTAACAAA
AGTACCATCAAATGCCTTTGAAGTACTTAAAAGTCTTAGAAGACTTTCTTTGTCTCATAATCC
TATTGAAGCAATACAGCCCTTTGCATTTAAAGGACTTGCCAATCTGGAATACCTCCTCCTGAA
AAATTCAAGAATTAGGAATGTTACTAGGGATGGGTTTAGTGGAATTAATAATCTTAAACATTT
GATCTTAAGTCATAATGATTTAGAGAATTTAAATTCTGACACATTCAGTTTGTTAAAGAATTT
AATTTACCTTAAGTTAGATAGAAACAGAATAATTAGCATTGATAATGATACATTTGAAAATAT
GGGAGCATCTTTGAAGATCCTTAATCTGTCATTTAATAATCTTACAGCCTTGCATCCAAGGGT
CCTTAAGCCGTTGTCTTCATTGATTCATCTTCAGGCAAATTCTAATCCTTGGGAATGTAACTG
CAAACTTTTGGGCCTTCGAGACTGGCTAGCATCTTCAGCCATTACTCTAAACATCTATTGTCA
GAATCCCCCATCCATGCGTGGCAGAGCATTACGTTATATTAACATTACAAATTGTGTTACATC
TTCAATAAATGTATCCAGAGCTTGGGCTGTTGTAAAATCTCCTCATATTCATCACAAGACTAC
TGCGCTAATGATGGCCTGGCATAAAGTAACCACAAATGGCAGTCCTCTGGAAAATACTGAGAC
TGAGAACATTACTTTCTGGGAACGAATTCCTACTTCACCTGCTGGTAGATTTTTTCAAGAGAA
TGCCTTTGGTAATCCATTAGAGACTACAGCAGTGTTACCTGTGCAAATACAACTTACTACTTC
TGTTACCTTGAACTTGGAAAAAAACAGTGCTCTACCGAATGATGCTGCTTCAATGTCAGGGAA
AACATCTCTAATTTGTACACAAGAAGTTGAGAAGTTGAATGAGGCTTTTGACATTTTGCTAGC
TTTTTTCATCTTAGCTTGTGTTTTAATCATTTTTTTGATCTACAAAGTTGTTCAGTTTAAACA
AAAACTAAAGGCATCAGAAAACTCAAGGGAAAATAGACTTGAATACTACAGCTTTTATCAGTC
AGCAAGGTATAATGTAACTGCCTCAATTTGTAACACTTCCCCAAATTCTCTAGAAAGTCCTGG
CTTGGAGCAGATTCGACTTCATAAACAAATTGTTCCTGAAAATGAGGCACAGGTCATTCTTTT
TGAACATTCTGCTTTA**TAA**CTCAACTAAATATTGTCTATAAGAAACTTCAGTGCCATGGACAT
GATTTAAACTGAAACCTCCTTATATAATTATATACTTTAGTTGGAAATATAATGAATTATATG
AGGTTAGCATTATTAAAATATGTTTTTTNTTAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 180

MCGLQFSLPCLRLFLVVTCYLLLLLHKEILGCSSVCQLCTGRQINCRNLGLSSIPKNFPESTV
FLYLTGNNISYINESELTGLHSLVALYLDNSNILYVYPKAFVQLRHLYFLFLNNNFIKRLDPG
IFKGLLNLRNLYLQYNQVSFVPRGVFNDLVSVQYLNLQRNRLTVLGSGTFVGMVALRILDLSN
NNILRISESGFQHLENLACLYLGSNNLTKVPSNAFEVLKSLRRLSLSHNPIEAIQPFAFKGLA
NLEYLLLKNSRIRNVTRDGFSGINNLKHLILSHNDLENLNSDTFSLLKNLIYLKLDRNRIISI
DNDTFENMGASLKILNLSFNNLTALHPRVLKPLSSLIHLQANSNPWECNCKLLGLRDWLASSA
ITLNIYCQNPPSMRGRALRYINITNCVTSSINVSRAWAVVKSPHIHHKTTALMMAWHKVTTNG
SPLENTETENITFWERIPTSPAGRFFQENAFGNPLETTAVLPVQIQLTTSVTLNLEKNSALPN
DAASMSGKTSLICTQEVEKLNEAFDILLAFFILACVLIIFLIYKVVQFKQKLKASENSRENRL
EYYSFYQSARYNVTASICNTSPNSLESPGLEQIRLHKQIVPENEAQVILFEHSAL

Signal peptide:

amino acids 1-41

Transmembrane domain:

amino acids 530-547

N-glycosylation sites.

amino acids 71-75, 76-80, 215-219, 266-270, 317-321, 331-335, 336-340, 400-404, 410-414, 451-455, 579-583

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 231-235

N-myristoylation sites.

amino acids 3-9, 69-75, 126-132, 174-180

ATP/GTP-binding site motif A (P-loop).

amino acids 506-514

FIGURE 181

GGCCTGGCGCGGCGCTCCGGTAAGGCGTGTGTGCGGCAGGGCGGGGACAGAACCGTCCTCTCG

GGCTCTGGGCGTGTCCGAGACCGCGCTCCCCGCCGAAATCAAGCTCCGAGTCATCCGTGTGGG

GCATTCGTCCCCCCTGGCACAGTTGGCCTCTTTCCAGAAGCCCGTTTTGTTTGTTTTACGTCT

AAATTCGCGTCGGTTCTTATTTCTCTCCCTGGCAAGGTCTGAAGACGGGTAGGAGAATAACCT

GTGTCAGCGTGTT**ATG**ATGCCGTCCCGTACCAACCTGGCTACTGGAATCCCCAGTAGTAAAGT

GAAATATTCAAGGCTCTCCAGCACAGACGATGGCTACATTGACCTTCAGTTTAAGAAAACCCC

TCCTAAGATCCCTTATAAGGCCATCGCACTTGCCACTGTGCTGTTTTTGATTGGCGCCTTTCT

CATTATTATAGGCTCCCTCCTGCTGTCAGGCTACATCAGCAAAGGGGGGGCAGACCGGGCCGT

TCCAGTGCTGATCATTGGCATTCTGGTGTTCCTACCCGGATTTTACCACCTGCGCATCGCTTA

CTATGCATCCAAAGGCTACCGTGGTTACTCCTATGATGACATTCCAGACTTTGATGAC**TAG**CA

CCCACCCCATAGCTGAGGAGGAGTCACAGTGGAACTGTCCCAGCTTTAAGATATCTAGCAGAA

ACTATAGCTGAGGACTAAGGAATTCTGCAGCTTGCAGATGTTTAAGAAAATAATGGCCAGATT

TTTTGGGTCCTTCCCAAAGATGTTAAGTGAACCTACAGTTAGCTAATTAGGACAAGCTCTATT

TTTCATCCCTGGGCCCTGACAAGTTTTTCCACAGGAATATGTATCATGGAAGAATAGAGGTTA

TTCTGTAATGGAAAGTGTTGCCTGCCACCACCCTCTGTAGAGCTGAGCATTTCTTTTAAATA

GTCTTCATTGCCAATTTGTTCTTGTAGCAAATGGAACAATGTGGTATGGCTAATTTCTTATTA

TTAAGTAGTTTATTTTAAAAATATCTGAGTATATTATCCTGTACACTTATCCCTACCTTCATG

TTCCAGTGGAAGACCTTAGTAAAATCAAAGATCAGTGAGTTCATCTGTAATATTTTTTTTACT

TGCTTTCTTACTGACAGCAACCAGGAATTTTTTTATCCTGCAGAGCAAGTTTTCAAAATGTAA

ATACTTCCTCTGTTTAACAGTCCTTGGACCATTCTGATCCAGTTCACCAGTAGGTTGGACAGC

ATATAATTTGCATCATTTTGTCCCTTGTAAATCAAGATGTTCTGCAGATTATTCCTTTAACGG

CCGGACTTTTGGCTGTTTCCTAATGAAACATGTAGTGGTTATTATTTAGAGTTTATAGCCGTA

TTGCTAGCACCTTGTAGTATGTCATCATTCTGCTCATGATTCCAAGGATCAGCCTGGATGCCT

AGAGGACTAGATCACCTTAGTTTGATTCTATTTTTTAGCTTGCAAAAAGTGACTTATATTCCA

AAGAAATTAAAATGTTGAAATCCAAATCCTAGAAATAAAATGAGTTTNNTTCCAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 182

MMPSRTNLATGIPSSKVKYSRLSSTDDGYIDLQFKKTPPKIPYKAIALATVLFLIGAFLIIIG

SLLLSGYISKGGADRAVPVLIIGILVFLPGFYHLRIAYYASKGYRGYSYDDIPDFDD

Transmembrane domains:

amino acids 45-66, 79-95

N-myristoylation sites.

amino acids 11-17, 75-81

FIGURE 183

```
CTAAAAAATACAAAAATTAGCTGGGCGTGGTGTCATGTACCTGTAATCCCAGCTACTCAAGAGGCTGAGGCAGGA
GAATCGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCAAGATTAAGTCACTGCACTCCAGCCTGGGTGACAGA
GCAAGACTCTGTATCAAAATAAATAAATAAAGTACAACTCTGGATGGGCATGGTGGCTTATGTCTGTAATCCCAG
CACTTTGGGAACTTGAGGCGGGTAGATTGCTTGAGTCCGGGAGTTTGAGACCAGTCTGGGTAATATGGTAACCCT
GTCTACCAAAAATACAGGTATTAGCCAGTCTCATAACTCGGTCTCAAAATAAATAAATACATACATACATAGATG
AAAATTTAAAAAATAAAGTCCAACTCAGCGGTTTTCAGCATATTTACAGAGTTGTACAATCTTCACCACTATCTA
ATTTCAGAACATTTTCATCACCCCCAAAAGAAACCTAACCCATTGACTATCTCTCCATTTCCTCCCTCTCCCTAG
CCTCTGGCAACCACTAATCTCTTTTTTGTCTCTATAGATTTGCCTATTTTGGACAGTTCATATACAAGGAATCAT
ACCACATGTAGCCTTTTGTGTCCGGCTTCTTTGATTAATAGAATGTTTTCAAGGCTCATCTATGCTGTAGCCTGT
ATCAGCACTTCATTCCTTTCTATGGCTGAATAATAGTCCACTGTAGGGATGTGCCATGTTTTTCCACTAGCTGAT
GGACATTTGGGTTGTTTCCACCTTCTGGCTATTATAAATATTGCTGCTATAAATATTCACTTACAAGTTTTTGTG
TGGACATATGTTTTTATTTCTTCTGGTATATCCTTCGGAGTGGAACTGCTGGATCAGGTGGTAACTCTAGGTCTA
ACCTGGCAGTTAAACAGAATCCTATGCATGCTGTAGTCCATGAGTTGAAATAAACACTTGACCCATAGTAAGTGC
CAGATCATCTTCATTTCACAGCAACCAGTAATTTCACAGATGAGGAAATGAAGGCTCCCAGAGGTGAACTGGCTT
TTCCCATTTGAGCAGTTCCAAGTCAGACAGTTAAAAAGTGGCAGGACCTGGAAGAGAAGCTAGTTCTTTCACCCT
GGCATTCAGGGCTGCCTCCTGGGCTACGGGGCTGGCATTTAGAATAGAGCTAAGGTCTGCTGCCAAGGCAGGTGC
CCCAGTCTGCCTCCTCTGTGTCCTTATTCCACTTTCTCTGCAGCCCTCCAGGGGACCCCTCTCTCAGCCACCCTC
TCTCTGGTGATGTCACAGTGCTGCCGGAAGATCAAAGATACGGTGCAGAAACTGGCTTCGGACCATAAGGACATT
CACAGCAGTGTATCCCGAGTGGGCAAAGCCATTGACAGGAACTTCGACTCTGAGATCTGTGGTGTTGTGTCAGAT
GCGGTGTGGGACGCGCGGGAACAGCAGCAGCAGATCCTGCAGATGGCCATCGTGGAACACCTGTATCAGCAGGGC
ATGCTCAGCGTGGCCGAGGAGCTGTGCCAGGAATCAACGCTGAATGTGGACTTGGATTTCAAGCAGCCTTTCCTA
GAGTTGAATCGAATCCTGGAAGCCCTGCACGAACAAGACCTGGGTCCTGCGTTGGAATGGGCCGTCTCCCACAGG
CAGCGCCTGCTGGAACTCAACAGCTCCCTGGAGTTCAAGCTGCACCGACTGCACTTCATCCGCCTCTTGGCAGGA
GGCCCCGCGAAGCAGCTGGAGGCCCTCAGCTATGCTCGGCACTTCCAGCCCTTTGCTCGGCTGCACCAGCGGGAG
ATCCAGGTGATGATGGGCAGCCTGGTGTACCTGCGGCTGGGCTTGGAGAAGTCACCCTACTGCCACCTGCTGGAC
AGCAGCCACTGGGCAGAGATCTGTGAGACCTTTACCCGGGACGCCTGTTCCCTGCTGGGGCTTTCTGTGGAGTCC
CCCCTTAGCGTCAGCTTTGCCTCTGGCTGTGTGGCGCTGCCTGTGTTGATGAACATCAAGGCTGTGATTGAGCAG
CGGCAGTGCACTGGGGTCTGGAATCACAAGGACGAGTTACCGATTGAGATTGAACTAGGCATGAAGTGCTGGTAC
GCTCATCTGTGGCCATGTTATCTCCCGAGATGCACTCAATAAGCTCATTAATGGAGGAAACACTCCGTGTTCGCT
TGCCCCATCCTCCGCCAGCAGACGTCAGATTCCAACCCTCCCATCAAGCTGAAGTGTCCCTACTGTCCATGGAG
CAGAACCCGGCAGATGGGAAACGCATCATATTCTGATTCCTACCTGGAAGGAATTTTGTTGAAAGGGGTTTTCAC
CTGTGAGCCTTGGTCTGTCTCGGTAGGGTGGTCAACTTCAGTGGACTGTGGTTGGTTTCAGAGCGCCTGGCTGAG
GAGTTCCACTGAGGGGAGCACTGGAGCAGCCCTTTGGCAGAGGCTGAGGAGGGAGATGGACCAGCCCACGCCTGG
CACCTGGCTCCATGGCATAAGGAAAGGGAGATGCTGGCCTCTGTGCTCCTGCTGTCTTTTCCTGTTTCTGTTTGC
GTTTGACTTAGTAGCAACCGACAGAGTGGCAAGGGATTTGGTCTTCAGCAGTAGACATCCTTCCACCCCTGCCCT
CAGCCAAGTCTCTTGCTGCCATGCCAATGCTATGTCCACCCTTGCCCCTCGGCCCAAGAGTGTCCAGCGGTGGCC
CACCTCTTCCTCCCACTACAGCCTCAACAGTATGTACCATCTCCCACTGTAAATAGTCCCAGTTAGAACGGAATG
CCGTTGTTTTATAACTTTGAACAAATGTATTTACTGCCCTTCTCAAAA
```

FIGURE 184

QCCRKIKDTVQKLASDHKDIHSSVSRVGKAIDRNFDSEICGVVSDAVWDAREQQQQILQMAIV

EHLYQQGMLSVAEELCQESTLNVDLDFKQPFLELNRILEALHEQDLGPALEWAVSHRQRLLEL

NSSLEFKLHRLHFIRLLAGGPAKQLEALSYARHFQPFARLHQREIQVMMGSLVYLRLGLEKSP

YCHLLDSSHWAEICETFTRDACSLLGLSVESPLSVSFASGCVALPVLMNIKAVIEQRQCTGVW

NHKDELPIEIELGMKCWYHSVFACPILRQQTSDSNPPIKLICGHVISRDALNKLINGGKLKCP

YCPMEQNPADGKRIIF

Transmembrane domain:

amino acids 222-241

N-glycosylation site.

amino acids 129-133

Tyrosine kinase phosphorylation site.

amino acids 151-159, 184-193

Amidation site.

amino acids 327-331

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 222-233

FIGURE 185

GAGCGACGCTGTCTCTAGTCGCTGATCCCAA**ATG**CACCGGCTCATCTTTGTCTACACTCTAAT

CTGCGCAAACTTTTGCAGCTGTCGGGACACTTCTGCAACCCCGCAGAGCGCATCCATCAAAGC

TTTGCGCAACGCCAACCTCAGGCGAGATGACTTGTACCGAAGAGATGAGACCATCCAGGTGAA

AGGAAACGGCTACGTGCAGAGTCCTAGATTCCCGAACAGCTACCCCAGGAACCTGCTCCTGAC

ATGGCGGCTTCACTCTCAGGAGAATACACGGATACAGCTAGTGTTTGACAATCAGTTTGGATT

AGAGGAAGCAGAAAATGATATCTGTAGGTATGATTTTGTGGAAGTTGAAGATATATCCGAAAC

CAGTACCATTATTAGAGGACGATGGTGTGGACACAAGGAAGTTCCTCCAAGGATAAAATCAAG

AACGAACCAAATTAAAATCACATTCAAGTCCGATGACTACTTTGTGGCTAAACCTGGATTCAA

GATTTATTATTCTTTGCTGGAAGATTTCCAACCCGCAGCAGCTTCAGAGACCAACTGGGAATC

TGTCACAAGCTCTATTTCAGGGGTATCCTATAACTCTCCATCAGTAACGGATCCCACTCTGAT

TGCGGATGCTCTGGACAAAAAAATTGCAGAATTTGATACAGTGGAAGATCTGCTCAAGTACTT

CAATCCAGAGTCATGGCAAGAAGATCTTGAGAATATGTATCTGGACACCCCTCGGTATCGAGG

CAGGTCATACCATGACCGGAAGTCAAAAGTTGACCTGGATAGGCTCAATGATGATGCCAAGCG

TTACAGTTGCACTCCCAGGAATTACTCGGTCAATATAAGAGAAGAGCTGAAGTTGGCCAATGT

GGTCTTCTTTCCACGTTGCCTCCTCGTGCAGCGCTGTGGAGGAAATTGTGGCTGTGGAACTGT

CAACTGGAGGTCCTGCACATGCAATTCAGGGAAAACCGTGAAAAAGTATCATGAGGTATTACA

GTTTGAGCCTGGCCACATCAAGAGGAGGGGTAGAGCTAAGACCATGGCTCTAGTTGACATCCA

GTTGGATCACCATGAACGATGCGATTGTATCTGCAGCTCAAGACCACCTCGA**TAA**GAGAATGT

GCACATCCTTACATTAAGCCTGAGAGAA

FIGURE 186

MHRLIFVYTLICANFCSCRDTSATPQSASIKALRNANLRRDDLYRRDETIQVKGNGYVQSPRF

PNSYPRNLLLTWRLHSQENTRIQLVFDNQFGLEEAENDICRYDFVEVEDISETSTIIRGRWCG

HKEVPPRIKSRTNQIKITFKSDDYFVAKPGFKIYYSLLEDFQPAAASETNWESVTSSISGVSY

NSPSVTDPTLIADALDKKIAEFDTVEDLLKYFNPESWQEDLENMYLDTPRYRGRSYHDRKSKV

DLDRLNDDAKRYSCTPRNYSVNIREELKLANVVFFPRCLLVQRCGGNCGCGTVNWRSCTCNSG

KTVKKYHEVLQFEPGHIKRRGRAKTMALVDIQLDHHERCDCICSSRPPR

Signal peptide:

amino acids 1-18

N-glycosylation site.

amino acids 270-274

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 262-266

Tyrosine kinase phosphorylation site.

amino acids 256-265

N-myristoylation sites.

amino acids 94-100, 186-192, 297-303, 298-304

TonB-dependent receptor proteins signature 1.

amino acids 1-56

FIGURE 187

C**ATG**CCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTG

TGTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGT

CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTT

CATATATGGGCAAAAGAAATGGCTGAATAAATCAGAATGCAGAAATATCAATAGAACCTACTG

TGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTG

GGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCTATCCTTTTTTAGAAACACAAAT

TGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC

AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCT

GAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAA

CCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTT

CGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGA

TCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGT

GTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACA

CCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAGATTCTTTGTGCCTGCTGA

AAAAATCGTGATTAACTTTATCACCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGA

TATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAA

CCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGA

AATTTTTTGTGACTCTGAAGAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAG

CAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACCACTGACAT

TTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATT

ATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCGCAAACGTTACAGTACTCATACACCCC

TCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGA

AGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAAACTGGCAGGCTGTGTATTCCTTCGCT

GTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGA

GGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAAC

CTATCTCATGCAATTCATGGAGGAATGGGGGTTATATGTGCAGATGGAAAAC**TGA**TGCCAACA

CTTCCTTTTGCCTTTTGTTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAA

GTACCTGGGATGAAGAAGTTTTTTCCAGTTTGTCAGTGTCTGTGAGAA

FIGURE 188

MPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGVKVTYTVQYF
IYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPFLETQI
GPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTWSQCVTN
HTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQSSEFKAKIIFWYVLPISITV
FLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAEKIVINFITLNISDDSKISHQD
MSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENTEGTSLTQQESLS
RTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLESQAALAVLGPQTLQYSYTP
QLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEE
GLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMEN

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 240-260

N-glycosylation sites.

amino acids 31-34, 72-75, 80-83, 171-174, 180-183, 189-192, 304-307, 523-526

Tyrosine kinase phosphorylation site.

amino acids 385-392, 518-526

N-myristoylation sites.

amino acids 53-58, 106-111, 368-373, 492-497

Tissue factor

amino acids 1-278

FIGURE 189

ATGTGCTGCTGGCCGCTGCTCCTGCTGTGGGGGCTGCTCCCCGGGACGGCGGCGGGGGGCTCG
GGCCGAACCTATCCGCACCGGACCCTCCTGGACTCGGAGGGCAAGTACTGGCTGGGCTGGAGC
CAGCGGGGGCAGCCAGATCGCCTTCCGCCTCCAGGTGCGCACTGCAGGCTACGTGGGCTTCGGC
TTCTCGCCCACCGGGGCCATGGCGTCCGCCGACATCGTCGTGGGCGGGGTGGCCCACGGGCGG
CCCTACCTCCAGGATTATTTTACAAATGCAAATAGAGAGTTGAAAAAAGATGCTCAGCAAGAT
TACCATCTAGAATATGCCATGGAAAATAGCACACACACAATAATTGAATTTACCAGAGAGCTG
CATACATGTGACATAAATGACAAGAGTATAACGGATAGCACTGTGAGAGTGATCTGGGCCTAC
CACCATGAAGATGCAGGAGAAGCTGGTCCCAAGTACCATGACTCCAATAGGGGCACCAAGAGT
TTGCGGTTATTGAATCCTGAGAAAACTAGTGTGCTATCTACAGCCTTACCATACTTTGATCTG
GTAAATCAGGACGTCCCCATCCCAAACAAAGATACAACATATTGGTGCCAAATGTTTAAGATT
CCTGTGTTCCAAGAAAAGCATCATGTAATAAAGGTTGAGCCAGTGATACAGAGAGGCCATGAG
AGTCTGGTGCACCACATCCTGCTCTATCAGTGCAGCAACAACTTTAACGACAGCGTTCTGGAG
TCCGGCCACGAGTGCTATCACCCCAACATGCCCGATGCATTCCTCACCTGTGAAACTGTGATT
TTTGCCTGGGCTATTGGTGGAGAGGGCTTTTCTTATCCACCTCATGTTGGATTATCCCTTGGC
ACTCCATTAGATCCGCATTATGTGCTCCTAGAAGTCCATTATGATAATCCCACTTATGAGGAA
GGCTTAATAGATAATTCTGGACTGAGGTTATTTTACACAATGGATATAAGGAAATATGATGCT
GGGGTGATTGAGGCTGGCCTCTGGGTGAGCCTCTTCCATACCATCCCTCCAGGGATGCCTGAG
TTCCAGTCTGAGGGTCACTGCACTTTGGAGTGCCTGGAAGAGGCTCTGGAAGCCGAAAAGCCA
AGTGGAATTCATGTGTTTGCTGTTCTTCTCCATGCTCACCTGGCTGGCAGAGGCATCAGGCTG
CGTCATTTTCGAAAAGGGAAGGAAATGAAATTACTTGCCTATGATGATGATTTTGACTTCAAT
TTCCAGGAGTTTCAGTATCTAAAGGAAGAACAAACAATCTTACCAGGAGATAACCTAATTACT
GAGTGTCGCTACAACACGAAAGATAGAGCTGAGATGACTTGGGGAGGACTAAGCACCAGGAGT
GAAATGTGTCTCTCATACCTTCTTTATTACCCAAGAATTAATCTTACTCGATGTGCAAGTATT
CCAGACATTATGGAACAACTTCAGTTCATTGGGGTTAAGGAGATCTACAGACCAGTCACGACC
TGGCCTTTCATTATCAAAAGTCCCAAGCAATATAAAAACCTTTCTTTCATGGATGCTATGAAT
AAGTTTAAATGGACTAAAAAGGAAGGTCTCTCCTTCAACAAGCTGGTCCTCAGCCTGCCAGTG
AATGTGAGATGTTCCAAGACAGACAATGCTGAGTGGTCGATTCAAGGAATGACAGCATTACCT
CCAGATATAGAAAGACCCTATAAAGCAGAACCTTTGGTGTGTGGCACGTCTTCTTCCTCTTCC
CTGCACAGAGATTTCTCCATCAACTTGCTTGTTTGCCTTCTGCTACTCAGCTGCACGCTGAGC
ACCAAGAGCTTG**TGA**TCAAAATTCTGTTGGACTTGACAATGTTTTCTATGATCTGAACCTGTC
ATTTGAAGTACAGGTTAAAGACTGTGTCCACTTTGGGCATGAAGAGTGTGGAGACTTTTCTTC
CCCATTTTCCCTCCCTCCTTTTTCCTTTCCATGTTACATGAGAGACATCAATCAGGTTCTCTT
CTCTTTCTTAGAAATACCTGATGTTATATATACATGGTCAATAAAATAAAACTGGCCTGACTT
AAGATAACCATTTTAAAAAATTGGGCTGTCATGTGGGAATAAAGAATTCTTTCTTTCCTAAA
AAAAAAAA

FIGURE 190

MCCWPLLLLWGLLPGTAAGGSGRTYPHRTLLDSEGKYWLGWSQRGSQIAFRLQVRTAGYVGFG
FSPTGAMASADIVVGGVAHGRPYLQDYFTNANRELKKDAQQDYHLEYAMENSTHTIIEFTREL
HTCDINDKSITDSTVRVIWAYHHEDAGEAGPKYHDSNRGTKSLRLLNPEKTSVLSTALPYFDL
VNQDVPIPNKDTTYWCQMFKIPVFQEKHHVIKVEPVIQRGHESLVHHILLYQCSNNFNDSVLE
SGHECYHPNMPDAFLTCETVIFAWAIGGEGFSYPPHVGLSLGTPLDPHYVLLEVHYDNPTYEE
GLIDNSGLRLFYTMDIRKYDAGVIEAGLWVSLFHTIPPGMPEFQSEGHCTLECLEEALEAEKP
SGIHVFAVLLHAHLAGRGIRLRHFRKGKEMKLLAYDDDFDFNFQEFQYLKEEQTILPGDNLIT
ECRYNTKDRAEMTWGGLSTRSEMCLSYLLYYPRINLTRCASIPDIMEQLQFIGVKEIYRPVTT
WPFIIKSPKQYKNLSFMDAMNKFKWTKKEGLSFNKLVLSLPVNVRCSKTDNAEWSIQGMTALP
PDIERPYKAEPLVCGTSSSSSLHRDFSINLLVCLLLLSCTLSTKSL

Signal peptide:
amino acids 1-18

Transmembrane domains:
amino acids 56-73, 378-393, 583-602

N-glycosylation sites.
amino acids 114-118, 247-251, 476-480, 517-521

N-myristoylation sites.
amino acids 11-17, 15-21, 20-26, 45-51, 68-74, 79-85, 290-296,
316-322, 337-343, 342-348, 456-462, 534-540, 582-588

Copper type II, ascorbate-dependent monooxygenases proteins.
amino acids 271-321, 422-474

FIGURE 191

GCTTCAGCTGAAGAAAGAGAGGA**ATG**AAGCGCCTTCTGCTTCTGTTTTTGTTCTTTATAACAT
TTTCTTCTGCATTTCCCTTAGTCCGGATGACGGAAAATGAAGAAAATATGCAACTGGCTCAGG
CATATCTCAACCAGTTCTACTCTCTTGAAATAGAAGGGAATCATCTTGTTCAAAGCAAGAATA
GGAGTCTCATAGATGACAAAATTCGGGAAATGCAAGCATTTTTTGGATTGACAGTGACTGGAA
AACTGGACTCAAACACCCTTGAGATCATGAAGACACCCAGGTGTGGGGTGCCTGATGTGGGCC
AGTATGGCTACACCCTCCCTGGGTGGAGAAAATACAACCTCACCTACAGAATAATAAACTATA
CTCCGGATATGGCACGAGCTGCTGTGGATGAGGCTATCCAAGAAGGTTTAGAAGTGTGGAGCA
AAGTCACTCCACTAAAATTCACCAAGATTTCAAAGGGGATTGCAGACATCATGATTGCCTTTA
GGACTCGAGTCCATGGTCGGTGTCCTCGCTATTTTGATGGTCCCTTGGGAGTGCTTGGCCATG
CCTTTCCTCCTGGTCCGGGTCTGGGTGGTGACACTCATTTTGATGAGGATGAAAACTGGACCA
AGGATGGAGCAGGATTCAACTTGTTTCTTGTGGCTGCTCATGAATTTGGTCATGCACTGGGGC
TCTCTCACTCCAATGATCAAACAGCCTTGATGTTCCCAAATTATGTCTCCCTGGATCCCAGAA
AATACCCACTTTCTCAGGATGATATCAATGGAATCCAGTCCATCTATGGAGGTCTGCCTAAGG
TACCTGCTAAGCCAAAGGAACCCACTATACCCCATGCCTGTGACCCTGACTTGACTTTTGACG
CTATCACAACTTTCCGCAGAGAAGTAATGTTCTTTAAAGGCAGGCACCTATGGAGGATCTATT
ATGATATCACGGATGTTGAGTTTGAATTAATTGCTTCATTCTGGCCATCTCTGCCAGCTGATC
TGCAAGCTGCATACGAGAACCCCAGAGATAAGATTCTGGTTTTTAAAGATGAAAACTTCTGGA
TGATCAGAGGATATGCTGTCTTGCCAGATTATCCCAAATCCATCCATACATTAGGTTTTCCAG
GACGTGTGAAGAAAATAGATGCAGCCGTCTGTGATAAGACCACAAGAAAAACCTACTTCTTTG
TGGGCATTTGGTGCTGGAGGTTTGATGAAATGACCCAAACCATGGACAAAGGATTCCCGCAGA
GAGTGGTAAAACACTTTCCTGGAATCAGTATCCGTGTTGATGCTGCTTTCCAGTACAAAGGAT
TCTTCTTTTTCAGCCGTGGATCAAAGCAATTTGAATACAACATTAAGACAAAGAATATTACCC
GAATCATGAGAACTAATACTTGGTTTCAATGCAAAGAACCAAAGAACTCCTCATTTGGTTTTG
ATATCAACAAGGAAAAAGCACATTCAGGAGGCATAAAGATATTGTATCATAAGAGTTTAAGCT
TGTTTATTTTTGGTATTGTTCATTTGCTGAAAAACACTTCTATTTATCAA**TAA**ATTCATAGAC
CTAAAATAAACCTCAACAGGTCTTTTAATATAAATTCTGCTTCAAAATAGAATAAAACCATTC
TTTAACAAC

FIGURE 192

MKRLLLLFLFFITFSSAFPLVRMTENEENMQLAQAYLNQFYSLEIEGNHLVQSKNRSLIDDKI

REMQAFFGLTVTGKLDSNTLEIMKTPRCGVPDVGQYGYTLPGWRKYNLTYRIINYTPDMARAA

VDEAIQEGLEVWSKVTPLKFTKISKGIADIMIAFRTRVHGRCPRYFDGPLGVLGHAFPPGPGL

GGDTHFDEDENWTKDGAGFNLFLVAAHEFGHALGLSHSNDQTALMFPNYVSLDPRKYPLSQDD

INGIQSIYGGLPKVPAKPKEPTIPHACDPDLTFDAITTFRREVMFFKGRHLWRIYYDITDVEF

ELIASFWPSLPADLQAAYENPRDKILVFKDENFWMIRGYAVLPDYPKSIHTLGFPGRVKKIDA

AVCDKTTRKTYFFVGIWCWRFDEMTQTMDKGFPQRVVKHFPGISIRVDAAFQYKGFFFFSRGS

KQFEYNIKTKNITRIMRTNTWFQCKEPKNSSFGFDINKEKAHSGGIKILYHKSLSLFIFGIVH

LLKNTSIYQ

Signal peptide:

amino acids 1-17

N-glycosylation sites.

amino acids 55-59, 110-114, 200-204, 452-456, 470-474, 508-512

N-myristoylation site.

amino acids 71-77, 205-211, 223-229

Hemopexin domain signature.

amino acids 171-202, 207-238, 318-334

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 213-223

Matrixins cysteine switch.

amino acids 89-97, 207-238

FIGURE 193

CACAATCAGGTCCCATTCTATAGATGGGGAAACTGAGGCTTGAGGTCACATAGGCGTCGTTCA

AGGCTGGTATACCTGCACCCTCTCCCATGTGAACAAC**ATG**GTTCTGGGTAATGGGGGCTGTCA

TCCAGTCTCCTCCCTGCCCCTGCTGGTGCACTTCCTGCCTCTGCTGGTGCACTTTCTGCCCCT

ACTGGTATATTTGCTGCCTCTGCTGGGGCGCTTCCTGCCTCGGCTGGTGTATCTCCTGCCCCT

GCTGGTGCACTTTCTGCCCCCGCTGATGCACTTCCTGCCTCTGCTGGTGCACTTCCTGGCTCT

GCTGGCACACTTCCTGCCTCTGCTGGTGCACTTCCTGGCTCTGCTGGCGCACTTTCCTGCCCC

TGCTGGTGTATTTCCTGCCCCTGCTGGTGTACTTCCTTCCCCTGCTGGTGCACTTCCTGCCTC

TGCTGGCGCACTTCTTGCCTCTCCAGGCCCTACC**TAG**CCTCTCCCTCTTATATATGGAAGTCT

TCCCAGTTCACTGACACTGGTAACAGGGACTCTGCTCTTGGTGTTGCTGTCTGCCCTGGGGAT

GGGCATCTGTGTCTTCCTTTACTACTGCTGGCTCAGGACCCAGAGCTTTGAAGCATGTCCAGA

TGCAGGTCCGGGCACCAGAGTCTAAGGAGCCCCTACACCCACCAGGATTTTCCAATAAAGAGA

TGTTCACCA

FIGURE 194

MVLGNGGCHPVSSLPLLVHFLPLLVHFLPLLVYLLPLLGRFLPRLVYLLPLLVHFLPPLMHFL

PLLVHFLALLAHFLPLLVHFLALLAHFPAPAGVFPAPAGVLPSPAGALPASAGALLASPGPT

Signal peptide:

amino acids 1-39

N-myristoylation sites.

amino acids 4-10, 109-115, 116-122

Leucine zipper pattern.

amino acids 14-36, 16-38, 17-39, 21-43, 24-46, 28-50, 31-53, 35-57, 38-60, 42-64, 45-67, 49-71, 52-74, 56-78, 59-81, 63-85, 65-87, 66-88

FIGURE 195

GGCAAGGCGGCGGCGGCGGCGGCGGCAGCCGCGGTGGCGGCGTGGGGAACATCTCGGCAGCCA
CCGCGCTTCTCCCGCTGGAGCGGGCGTCCAGCTTGGCTGCCCTCGGTCCTTCCCTGCCACGTT
TCGGGTCGCCCTGCACCCCCCACCCAGGCTCGCTTCTCTTCGAAGCGGGAAGGGCGCCTTGCA
GGATCCTGCCGCCCCTCCAACCGGATCCTGGGTCTAGAGCTCCCCAGAGCGAGGCGCTCGCCA
GGACTCCTGCCCCGCCAACCCTGACCGCCGGGGGGTGCCCCGGGACGTAGCGCCGCGGAGAG
GAAGCGGCAAAGGGGACC**ATG**CGGCGCCTGACTCGTCGGCTGGTTCTGCCAGTCTTCGGGGTG
CTCTGGATCACGGTGCTGCTGTTCTTCTGGGTAACCAAGAGGAAGTTGGAGGTGCCGACGGGA
CCTGAAGTGCAGACCCCTAAGCCTTCGGACGCTGACTGGGACGACCTGTGGGACCAGTTTGAT
GAGCGGCGGTATCTGAATGCCAAAAAGTGGCGCGTTGGTGACGACCCCTATAAGCTGTATGCT
TTCAACCAGCGGGAGAGTGAGCGGATCTCCAGCAATCGGGCCATCCCGGACACTCGCCATCTG
AGATGCACACTGCTGGTGTATTGCACGGACCTTCCACCCACTAGCATCATCATCACCTTCCAC
AACGAGGCCCGCTCCACGCTGCTCAGGACCATCCGCAGTGTATTAAACCGCACCCCTACGCAT
CTGATCCGGGAAATCATATTAGTGGATGACTTCAGCAATGACCCTGATGACTGTAAACAGCTC
ATCAAGTTGCCCAAGGTGAAATGCTTGCGCAATAATGAACGGCAAGGTCTGGTCCGGTCCCGG
ATTCGGGGCGCTGACATCGCCCAGGGCACCACTCTGACTTTCCTCGACAGCCACTGTGAGGTG
AACAGGGACTGGCTCCAGCCTCTGTTGCACAGGGTCAAAGAGGACTACACGCGGGTGGTGTGC
CCTGTGATCGATATCATTAACCTGGACACCTTCACCTACATCGAGTCTGCCTCGGAGCTCAGA
GGGGGGTTTGACTGGAGCCTCCACTTCCAGTGGGAGCAGCTCTCCCCAGAGCAGAAGGCTCGG
CGCCTGGACCCCACGGAGCCCATCAGGACTCCTATCATAGCTGGAGGGCTCTTCGTGATCGAC
AAAGCTTGGTTTGATTACCTGGGGAAATATGATATGGACATGGACATCTGGGGTGGGGAGAAC
TTTGAAATCTCCTTCCGAGTGTGGATGTGCGGGGGCAGCCTAGAGATCGTCCCCTGCAGCCGA
GTGGGGCACGTCTTCCGGAAGAAGCACCCCTACGTTTTCCCTGATGGAAATGCCAACACGTAT
ATAAAGAACACCAAGCGGACAGCTGAAGTGTGGATGGATGAATACAAGCAATACTATTACGCT
GCCCGGCCATTCGCCCTGGAGAGGCCCTTCGGGAATGTTGAGAGCAGATTGGACCTGAGGAAG
AATCTGCGCTGCCAGAGCTTCAAGTGGTACCTGGAGAATATCTACCCTGAACTCAGCATCCCC
AAGGAGTCCTCCATCCAGAAGGGCAATATCCGACAGAGACAGAAGTGCCTGGAATCTCAAAGG
CAGAACAACCAAGAAACCCCAAACCTAAAGTTGAGCCCCTGTGCCAAGGTCAAAGGCGAAGAT
GCAAAGTCCCAGGTATGGGCCTTCACATACACCCAGCAGATCCTCCAGGAGGAGCTGTGCCTG
TCAGTCATCACCTTGTTCCCTGGCGCCCCAGTGGTTCTTGTCCTTTGCAAGAATGGAGATGAC
CGACAGCAATGGACCAAAACTGGTTCCCACATCGAGCACATAGCATCCCACCTCTGCCTCGAT
ACAGATATGTTCGGTGATGGCACCGAGAACGGCAAGGAAATCGTCGTCAACCCATGTGAGTCC
TCACTCATGAGCCAGCACTGGGACATGGTGAGCTCT**TGA**GGACCCCTGCCAGAAGCAGCAAGG
GCCATGGGGTGGTGCTTCCCTGGACCAGAACAGACTGGAAACTGGGCAGCAAGCAGCCTGCAA
CCACCTCAGACATCCTGGACTGGGAGGTGGAGGCAGAGCCCCCCAGGACAGGAGCAACTGTCT
CAGGGAGGACAGAGGAAAACATCACAAGCCAATGGGCTCAAAGACAAATCCCACATGTTCTCA
AGGCCGTTAAGTTCCAGTCCTGGCCAGTCATTCCCTGATTGGTATCTGGAGACAGAAACCTAA
TGGGAAGTGTTTATTGTTCCTTTTCCTACAAAGGAAGCAGTCTCTGGAGGCCAGAAAGAAAG
CCTTCTTTTTCACTAGGCCAGGACTACATTGAGAGATGAAGAATGGAGGTTGTTTCCAAAAGA
AATAAAGAGAAACTTAGAAGTTGTCTCTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAA

FIGURE 196

MRRLTRRLVLPVFGVLWITVLLFFWVTKRKLEVPTGPEVQTPKPSDADWDDLWDQFDERRYLN

AKKWRVGDDPYKLYAFNQRESERISSNRAIPDTRHLRCTLLVYCTDLPPTSIIITFHNEARST

LLRTIRSVLNRTPTHLIREIILVDDFSNDPDDCKQLIKLPKVKCLRNNERQGLVRSRIRGADI

AQGTTLTFLDSHCEVNRDWLQPLLHRVKEDYTRVVCPVIDIINLDTFTYIESASELRGGFDWS

LHFQWEQLSPEQKARRLDPTEPIRTPIIAGGLFVIDKAWFDYLGKYDMDMDIWGGENFEISFR

VWMCGGSLEIVPCSRVGHVFRKKHPYVFPDGNANTYIKNTKRTAEVWMDEYKQYYYAARPFAL

ERPFGNVESRLDLRKNLRCQSFKWYLENIYPELSIPKESSIQKGNIRQRQKCLESQRQNNQET

PNLKLSPCAKVKGEDAKSQVWAFTYTQQILQEELCLSVITLFPGAPVVLVLCKNGDDRQQWTK

TGSHIEHIASHLCLDTDMFGDGTENGKEIVVNPCESSLMSQHWDMVSS

Transmembrane domain:

amino acids 475-493

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 2-6

Tyrosine kinase phosphorylation sites.

amino acids 68-75, 401-409

N-myristoylation sites.

amino acids 178-184, 186-192, 192-198, 346-352, 383-389, 526-532

FIGURE 197

```
GCAGCTCACCCTTCGCAGCCGCGATGGGGGAAGACGACGCCGCGCTTCGGGCTGGCAGCAGGGGGCTCTCCGACC
CGTGGGCAGACTCAGTGGGAGTGCGACCCCGCACCACGGAGCGCCACATCGCCGTACACAAGCGGCTTGTGCTGG
CCTTCGCTGTGTCCCTCGTGGCATTGCTCGCGGTCACAATGCTCGCTGTGCTGCTCAGCCTGCGCTTCGACGAGT
GCGGGGCGAGTGCCACGCCAGGCGCCGACGGTGGCCCCTCAGGCTTTCCGGAGCGCGGCGGCAACGGGAGCCTCC
CTGGATCGGCCCGGCGCAACCACCACGCAGGCGGGGACTCCTGGCAGCCCGAGGCGGGTGGGGTGGCCAGTCCGG
GGACCACGTCGGCCCAGCCGCCGTCGGAGGAGGAGCGGGAGCCGTGGGAGCCGTGGACGCAGCTGCGCCTGTCGG
GCCACCTGAAGCCGCTGCACTACAATCTGATGCTCACCGCCTTCATGGAGAACTTCACCTTCTCCGGGGAGGTCA
ACGTGGAGATCGCGTGCCGGAACGCCACCCGCTACGTAGTGCTGCACGCTTCCCGAGTGGCGGTGGAGAAAGTGC
AGCTGGCCGAGGACCGGGCGTTCGGGGCTGTCCCTGTAGCCGGTTTTTTCCTCTACCCGCAAACCCAGGTCTTAG
TGGTGGTGCTGAATAGGACACTGGACGCGCAGAGGAATTACAATCTGAAGATTATCTACAACGCGCTCATCGAGA
ATGAGCTCCTGGGCTTCTTCCGCAGCTCCTATGTGCTCCACGGGGAGAGAAGATTCCTTGGTGTTACTCAGTTTT
CGCCTACACATGCCAGAAAGGCATTTCCTTGTTTTGATGAGCCAATCTACAAGGCTACTTTCAAAATCAGCATCA
AGCATCAAGCAACCTATTTATCTTTATCTAATATGCCAGTGGAAACTTCCGTGTTTGAGGAAGATGGATGGGTTA
CGGATCACTTTTCACAGACCCCTCTCATGTCCACATATTATTTAGCCTGGGCAATTTGCAACTTCACATACAGAG
AAACTACCACCAAGAGTGGGGTTGTAGTACGATTATATGCAAGACCTGATGCTATCAGAAGAGGATCCGGGGACT
ATGCTCTCCATATAAACAAAGAGATTAATAGAATTTTATGAAGACTACTTTAAAGTGCCCTATTCCTTGCCAAAAC
TAGATCTTTTAGCTGTGCCTAAGCATCCGTATGCTGCTATGGAGAACTGGGGACTAAGTATTTTTGTGGAACAAA
GAATACTGCTGGATCCCAGTGTTTCATCTATTTCTTATTTGCTGGATGTCACCATGGTCATTGTTCATGAGATAT
GTCACCAGTGGTTTGGTGACCTTGTGACGCCTGTGTGGTGGGAAGACGTGTGGCTGAAGGAAGGGTTTGCTCACT
ACTTTGAATTTGTTGGTACAGACTACCTCTATCCTGGCTGGAACATGGAAAAGCAGAGGTTTCTGACCGATGTTC
TGCATGAAGTGATGCTGCTGGACGGTTTGGCCAGTTCCCATCCAGTATCACAGGAAGTGCTGCAGGCAACAGATA
TTGACAGGGTGTTTGACTGGATCGCATATAAAAAGGGTGCTGCTTTAATAAGAATGCTGGCTAATTTTATGGGCC
ATTCAGTTTTCCAGAGGGGGTTTGCAAGATTATTTAACCATTCATAAGTATGGTAATGCAGCCAGAAATGATCTCT
GGAATACATTATCGGAGGCTTTAAAAAGAAATGGGAAATATGTAAATACAAGAAGTAATGGATCAGTGGACAC
TCCAGATGGGTTATCCTGTTATCACCATCTTGGGAAACACAACAGCAGAAAATAGAATAATAATTACCCAACAGC
ATTTTATCTATGATATCAGTGCTAAAACTAAAGCACTTAAACTTCAGAATAACAGTTACCTGTGGCAGATTCCAT
TAACTATTGTGGTAGGAAATAGAAGCCATGTGTCTTCAGAAGCAATTATTTGGGTGTCTAACAAATCAGAGCACC
ACAGAATAACTTATTTGGACAAAGGAAGCTGGCTGCTGGGGAACATCAATCAAACTGGCTATTTTAGAGTCAACT
ATGACCTAAGGAACTGGAGATTATTAATTGATCAATTAATCCGGAATCATGAGGTTCTTTCTGTCAGTAACCGAG
CGGGCTTGATCGATGATGCCTTCAGCCTAGCCAGGGCTGGCTATTTGCCTCAGAATATTCCTCTGGAGATTATCA
GATACCTGTCTGAGGAGAAGGATTTTCTTCCTTGGCATGCTGCCAGCCGAGCTCTTTATCCTCTAGATAAATTAC
TGGACCGCATGGAAAACTACAACATTTTCAATGAATATATTTTAAAGCAAGTTGCAACAACATATATCAAGCTTG
GGTGGCCGAAAAATAATTTTAATGGATCTCTTGTTCAAGCATCCTACCAACATGAAGAACTACGTAGAGAAGTTA
TAATGCTGGCCTGCAGTTTTGGCAACAAGCACTGTCACCAACAGGCATCAACACTTATTTCAGATTGGATTTCCA
GCAACAGGAACAGAATACCACTAAATGTTAGAGACATCGTATACTGTACAGGAGTGTCACTACTGGATGAGGATG
TCTGGGAATTCATATGGATGAAATTCCATTCCACCACAGCAGTTTCTGAGAAGAAAATATTATTGGAAGCCTTAA
CTTGCAGTGATGACAGGAATTTATTAAACAGGCTTCTAAATCTGTCACTGAATTCTGAGGTGGTGCTGGATCAAG
ATGCAATTGATGTCATAATCCATGTAGCTCGAAATCCACATGGTCGAGACCTTGCCTGGAAGTTTTTCAGGGATA
AATGGAAGATATTAAATACCAGGTATGGAGAAGCATTGTTTATGTATTCCAAACTCATCAGTGGTGTCACAGAAT
TTCTTAATACTGAAGGTGAACTCAAAGAGCTCAAGAACTTCATGAAAAACTATGATGGGGTAGCTGCTGCTTCTT
TCTCACGAGCTGTGGAAACTGTCGAAGCCAATGTGCGCTGGAAAATGCTTTACCAAGACGAGCTTTTCCAATGGT
TAGGAAAAGCTCTAAGACACTAATATATGTATCTTATAAACAAACAATTCAACTCAGAAGTTTATGAGAAGACAC
GCTTTTTGTGGAATGAGGAAAATGTACTACCTAGAAAATGGCCAGATTTTCAGTGTTAACGTGTGGGAGGAATTT
TTTTTTTTAGTTTTTATTTTTTGGTTTTGGGGGATATTTTTTATTTGTTTCATTCATTCTGTTCTGTTTCTCTAC
TGGGTGTTCCTCTCTAAAGAAACTCTTGCAAGTGAAACTAGCCATGATTGCTTCAGCTGTACATTCCTTGCTGTA
CAGGACCAAATATGATAGTGATGCATGTTGATGTTACAGTCAATTTGGAAAAACATATTCAGAATATCTGTGCAT
GGATATATTGTCCTGCCTGTGTTCCAGCATGCTTATTTCAAACGTCCAGTGTTGTGTGTGAATATGTGTTACACC
TAGGATGGGCATTATGCAAAAGCACAAAGATTATATATGACAATCAGTATTGCAATGAAAGAAAAACTAAAAACA
GAAATGATATTCTCAATTTTGGGCAATGTGAGAGGTAAAATAGCCCTTGACATGATGAACATCACTTATTTCAGC
ACTTGGATTGTCTGGCAATGATTACTGTGTTGCTAACTCATTTTCTTTGAGTTAAAGCTGTGTATACATTTTAAA
AGGCATATAGATAGTGTATGCATATGTATATGTACATAGGGAAGCCCCATATGTATATAGTATGTTGTACACTGC
ACATGTACAAAGAATGTCTTCAGATCAAAGAAAATTTATCTCTTTTTATAAACTTAAGGACAGTTGCAAAAGGCT
TCAAGGAATTTATCTCAACATTATTCTTTCTATGTCCTAACTAAATTTCTCAACTGTTATGAATTTTTTCATCTAC
TTCTTGAACAGTGGTCTATTCTGCTACATGAAGATGAATACAAACAAAATTTTTGTATAAACTCCCAAAAAAAAA
AAAAAAAAAA
```

FIGURE 198

MGEDDAALRAGSRGLSDPWADSVGVRPRTTERHIAVHKRLVLAFAVSLVALLAVTMLAVLLSL

RFDECGASATPGADGGPSGFPERGGNGSLPGSARRNHHAGGDSWQPEAGGVASPGTTSAQPPS

EEEREPWEPWTQLRLSGHLKPLHYNLMLTAFMENFTFSGEVNVEIACRNATRYVVLHASRVAV

EKVQLAEDRAFGAVPVAGFFLYPQTQVLVVVLNRTLDAQRNYNLKIIYNALIENELLGFFRSS

YVLHGERRFLGVTQFSPTHARKAFPCFDEPIYKATFKISIKHQATYLSLSNMPVETSVFEEDG

WVTDHFSQTPLMSTYYLAWAICNFTYRETTTKSGVVVRLYARPDAIRRGSGDYALHITKRLIE

FYEDYFKVPYSLPKLDLLAVPKHPYAAMENWGLSIFVEQRILLDPSVSSISYLLDVTMVIVHE

ICHQWFGDLVTPVWWEDVWLKEGFAHYFEFVGTDYLYPGWNMEKQRFLTDVLHEVMLLDGLAS

SHPVSQEVLQATDIDRVFDWIAYKKGAALIRMLANFMGHSVFQRGLQDYLTIHKYGNAARNDL

WNTLSEALKRNGKYVNIQEVMDQWTLQMGYPVITILGNTTAENRIIITQQHFIYDISAKTKAL

KLQNNSYLWQIPLTIVVGNRSHVSSEAIIWVSNKSEHHRITYLDKGSWLLGNINQTGYFRVNY

DLRNWRLLIDQLIRNHEVLSVSNRAGLIDDAFSLARAGYLPQNIPLEIIRYLSEEKDFLPWHA

ASRALYPLDKLLDRMENYNIFNEYILKQVATTYIKLGWPKNNFNGSLVQASYQHEELRREVIM

LACSFGNKHCHQQASTLISDWISSNRNRIPLNVRDIVYCTGVSLLDEDVWEFIWMKFHSTTAV

SEKKILLEALTCSDDRNLLNRLLNLSLNSEVVLDQDAIDVIIHVARNPHGRDLAWKFFRDKWK

ILNTRYGEALFMYSKLISGVTEFLNTEGELKELKNFMKNYDGVAAASFSRAVETVEANVRWKM

LYQDELFQWLGKALRH

Transmembrane domain:

amino acids 44-63

N-glycosylation sites.

amino acids 89-93, 160-164, 175-179, 222-226, 338-342, 605-609,
634-638, 649-653, 663-667, 684-688, 800-804, 906-910

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 362-366

Tyrosine kinase phosphorylation site.

amino acids 520-528

N-myristoylation sites.

amino acids 78-84, 87-93, 90-96, 118-124, 501-507, 604-610,
825-831, 987-993

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 437-447

FIGURE 199

GCGCCCGGCGCAGCTCGGCCAGAGCGACCGCGGGGCTGAGCGCGCGTCCGCCCAGGGGGCTCCGGAAGCTGCCCC
GGCCCGCGGCCTCCTCCCTCGCTCCCGCTTCCCCTTTCTCGCTCACCGCCGCCCTCCTTCCCCAGCTCCCTCGCC
GTCCGCCCGCCCCACAGCCAGCGGCTCCGCGCCCCCTGCAGCCACG**ATG**CCCGCGGCCCGGCCGCCCGCCGCGGG
ACTCCGCGGGATCTCGCTGTTCCTCGCTCTGCTCCTGGGGAGCCCGGCGGCAGCGCTGGAGCGAGATGCTCTTCC
CGAGGGAGATGCTAGCCCTTTGGGTCCTTACCTCCTGCCCTCAGGAGCCCCGGAGAGAGGCAGTCCTGGCAAAGA
GCACCCTGAAGAGAGAGTGGTAACAGCGCCCCCCAGTTCCTCACAGTCGGCGGAAGTGCTGGGCGAGCTGGTGCT
GGATGGGACCGCACCCTCTGCACATCACGACATCCCAGCCCTGTCACCGCTGCTTCCAGAGGAGGCCCGCCCCAA
GCACGCCTTGCCCCCCAAGAAGAAACTGCCTTCGCTCAAGCAGGTGAACTCTGCCAGGAAGCAGCTGAGGCCCAA
GGCCACCTCCGCAGCCACTGTCCAAAGGGCAGGGTCCCAGCCAGCGTCCCAGGGCCTAGATCTCCTCTCCTCCTC
CACGGAGAAGCCTGGCCCACCGGGGGACCCGGACCCCATCGTGGCCTCCGAGGAGGCATCAGAAGTGCCCCTTTG
GCTGGATCGAAAGGAGAGTGCGGTCCCTACAACACCCGCACCCCTGCAAATCTCCCCCTTCACTTCGCAGCCCTA
TGTGGCCCACACACTCCCCCAGAGGCCAGAACCCGGGGAGCCTGGGCCTGACATGGCCCAGGAGGCCCCCCAGGA
GGACACCAGCCCCATGGCCCTGATGGACAAAGGTGAGAATGAGCTGACTGGGTCAGCCTCAGAGGAGAGCCAGGA
GACCACTACCTCCACCATTATCACCACCACGGTCATCACCACCGAGCAAGCACCAGCTCTCTGCAGTGTGAGCTT
CTCCAATCCTGAGGGGTACATTGACTCCAGCGACTACCCCACTGCTGCCCCTCAACAACTTTCTGGAGTGCACATA
CAACGTGACAGTCTACACTGGCTATGGGGTGGAGCTCCAGGTGAAGAGTGTGAACCTGTCCGATGGGGAACTGCT
CTCCATCCGCGGGGTGGACGGCCCTACCCTGACCGTCCTGGCCAACCAGACACTCCTGGTGGAGGGGCAGGTAAT
CCGAAGCCCCACCAACACCATCTCCGTCTACTTCCGGACCTTCCAGGACGACGGCCTTGGGACCTTCCAGCTTCA
CTACCAGGCCTTCATGCTGAGCTGCAACTTTCCCCGCCGGCCTGACTCTGGGGATGTCACGGTGATGGACCTGCA
CTCAGGTGGGGTGGCCCACTTTCACTGCCACCTGGGCTATGAGCTCCAGGGCGCTAAGATGCTGACATGCATCAA
TGCCTCCAAGCCGCACTGGAGCAGCCAGGAGCCCATCTGCTCAGCTCCTTGTGGAGGGGCAGTGCACAATGCCAC
CATCGGCCGCGTCCTCTCCCCAAGTTACCCTGAAAACACAAATGGGAGCCAATTCTGCATCTGGACGATTGAAGC
TCCAGAGGGCCAGAAGCTGCACCTGCACTTTGAGAGGCTGTTGCTGCATGACAAGGACAGGATGACGGTTCACAG
CGGGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCTTTTGAGGGCCTGCTGAG
CGAAGGCAACACCATCCGCATCGAGTTCACGTCCGACCAGGCCCGGGCGGCCTCCACCTTCAACATCCGATTTGA
AGCGTTTGAGAAAGGCCACTGCTATGAGCCCTACATCCAGAATGGGAACTTCACTACATCCGACCCGACCTATAA
CATTGGGACTATAGTGGAGTTCACCTGCGACCCCGGCCACTCCCTGGAGCAGGGCCCGGCCATCATCGAATGCAT
CAATGTGCGGGACCCATACTGGAATGACACAGAGCCCCTGTGCAGAGCCATGTGTGGTGGGGAGCTCTCTGCTGT
GGCTGGGGTGGTATTGTCCCCAAACTGGCCCGAGCCCTACGTGGAAGGTGAAGATTGTATCTGGAAGATCCACGT
GGGAGAAGAGAAACGGATCTTCTTAGATATCCAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGA
TGGCGACGAGGTCATGCCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCCCAGAAACTGTACTCCTCCAC
GCCAGACTTAACCATCCAGTTCCATTCGGGACCCTGCTGGCCTCATCTTTGGAAAGGGCCAGGGATTTATCATGAA
CTACATAGAGGTATCAAGGAATGACTCCTGCTCGGATTTACCCGAGATCCAGAATGGCTGGAAAACCACTTCTCA
CACGGAGTTGGTGCGGGGAGCCAGAATCACCTACCAGTGTGACCCCGGCTATGACATCGTGGGGAGTGACACCCT
CACCTGCCAGTGGGACCTCAGCTGGAGCAGCGACCCCCCATTTTGTGAGAAAATTATGTACTGCACCGACCCCGG
AGAGGTGGATCACTCGACCCGCTTAATTTCGGATCCTGTGCTGCTGGTGGGGACCACCATCCAATACACCTGCAA
CCCCGGTTTTGTGCTTGAAGGGAGTTCTCTTCTGACCTGCTACAGCCGTGAAACAGGGACTCCCATCTGGACGTC
TCGCCTGCCCCACTGCGTTTCGGAGGAGTCCCTGGCATGTGACAACCCAGGGCTGCCTGAAAATGGATACCAAAT
CCTGTACAAGCGACTCTACCTGCCAGGAGAGTCCCTCACCTTCATGTGCTACGAAGGCTTTGAGCTCATGGGTGA
AGTGACCATCCGCTGCATCCTGGGACAGCCATCCCACTGGAACGGGCCCCTGCCCGTGTGTAAAGTTAATCAAGA
CAGTTTTGAACATGCTTTAGAAGCAGAAGCGGCAGCAGAGACGTCGCTGGAAGGGGGGAACATGGCCCTGGCTAT
CTTCATCCCGGTCCTCATCATCTCCTTACTGCTGGGAGGAGCCTACATTTACATCACAAGATGTCGCTACTATTC
CAACCTCCGCCTGCCTCTGATGTACTCCCACCCCTACAGCCAGATCACCGTGGAAACCGAGTTTGACAACCCCAT
TTACGAGACAGGGGAAACCAGAGAGTATGAGGTTTCTATC**TAA**AGAGAGCTACACTTGAGAAGGGGACTTGTGAA
CTCAACCACAATCTCCTCGAGACATTCATCCAGAGACCATGTGGCACTTGATTGAAACCCCAGAATGTCGACTGT
CTTTTGTTTAGACTCTTTATCAAAGGTTTACTGTTTTCTTCCCTGTATTTATTATATTTAAAAGTGAAAAAAAAA
AAAAAAAAAA

FIGURE 200

MPAARPPAAGLRGISLFLALLLGSPAAALERDALPEGDASPLGPYLLPSGAPERGSPGKEHPE
ERVVTAPPSSSQSAEVLGELVLDGTAPSAHHDIPALSPLLPEEARPKHALPPKKKLPSLKQVN
SARKQLRPKATSAATVQRAGSQPASQGLDLLSSSTEKPGPPGDPDPIVASEEASEVPLWLDRK
ESAVPTTPAPLQISPFTSQPYVAHTLPQRPEPGEPGPDMAQEAPQEDTSPMALMDKGENELTG
SASEESQETTTSTIITTTVITTEQAPALCSVSFSNPEGYIDSSDYPLLPLNNFLECTYNVTVY
TGYGVELQVKSVNLSDGELLSIRGVDGPTLTVLANQTLLVEGQVIRSPTNTISVYFRTFQDDG
LGTFQLHYQAFMLSCNFPRRPDSGDVTVMDLHSGGVAHFHCHLGYELQGAKMLTCINASKPHW
SSQEPICSAPCGGAVHNATIGRVLSPSYPENTNGSQFCIWTIEAPEGQKLHLHFERLLLHDKD
RMTVHSGQTNKSALLYDSLQTESVPFEGLLSEGNTIRIEFTSDQARAASTFNIRFEAFEKGHC
YEPYIQNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPAIIECINVRDPYWNDTEPLCRAMCGG
ELSAVAGVVLSPNWPEPYVEGEDCIWKIHVGEEKRIFLDIQFLNLSNSDILTIYDGDEVMPHI
LGQYLGNSGPQKLYSSTPDLTIQFHSDPAGLIFGKGQGFIMNYIEVSRNDSCSDLPEIQNGWK
TTSHTELVRGARITYQCDPGYDIVGSDTLTCQWDLSWSSDPPFCEKIMYCTDPGEVDHSTRLI
SDPVLLVGTTIQYTCNPGFVLEGSSLLTCYSRETGTPIWTSRLPHCVSEESLACDNPGLPENG
YQILYKRLYLPGESLTFMCYEGFELMGEVTIRCILGQPSHWNGPLPVCKVNQDSFEHALEAEA
AAETSLEGGNMALAIFIPVLIISLLLGGAYIYITRCRYYSNLRLPLMYSHPYSQITVETEFDN
PIYETGETREYEVSI

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 893-915

N-glycosylation sites.

amino acids 311-315, 328-332, 350-354, 435-439, 458-462, 474-478, 514-518, 576-580, 618-622, 674-678, 742-746

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 188-192

N-myristoylation sites.

amino acids 23-29, 87-93, 146-152, 454-460, 475-481, 575-581, 629-635, 695-701, 723-729, 766-772, 877-883, 953-959

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 383-394

FIGURE 201

G**ATG**GCTACGGCAGGGGGTGGCTCTGGGGCTGACCCGGGAAGTCGGGGTCTCCTTCGCCTTCT
GTCTTTCTGCGTCCTACTAGCAGGTTTGTGCAGGGGAAACTCAGTGGAGAGGAAGATATATAT
CCCCTTAAATAAAACAGCTCCCTGTGTTCGCCTGCTCAACGCCACTCATCAGATTGGCTGCCA
GTCTTCAATTAGTGGAGACACAGGGGTTATCCACGTAGTAGAGAAAGAGGAGGACCTACAGTG
GGTATTGACTGATGGCCCCAACCCCCCTTACATGGTTCTGCTGGAGAGCAAGCATTTTACCAG
GGATTTAATGGAGAAGCTGAAAGGGAGAACCAGCCGAATTGCTGGTCTTGCAGTGTCCTTGAC
CAAGCCCAGTCCTGCCTCAGGCTTCTCTCCTAGTGTACAGTGCCCAAATGATGGGTTTGGTGT
TTACTCCAATTCCTATGGGCCAGAGTTTGCTCACTGCAGAGAAATACAGTGGAATTCGCTGGG
CAATGGTTTGGCTTATGAAGACTTTAGTTTCCCCATCTTTCTTCTTGAAGATGAAAATGAAAC
CAAAGTCATCAAGCAGTGCTATCAAGATCACAACCTGAGTCAGAATGGCTCAGCACCAACCTT
CCCACTATGTGCCATGCAGCTCTTTTCACACATGCATGCTGTCATCAGCACTGCCACCTGCAT
GCGGCGCAGCTCCATCCAAAGCACCTTCAGCATCAACCCAGAAATCGTCTGTGACCCCCTGTC
TGATTACAATGTGTGGAGCATGCTAAAGCCTATAAATACAACTGGGACATTAAAGCCTGACGA
CAGGGTTGTGGTTGCTGCCACCCGGCTGGATAGTCGTTCCTTTTTCTGGAATGTGGCCCCAGG
GGCTGAAAGCGCAGTGGCTTCCTTTGTCACCCAGCTGGCTGCTGCTGAAGCTTTGCAAAAGGC
ACCTGATGTGACCACCCTGCCCCGCAATGTCATGTTTGTCTTCTTTCAAGGGGAAACTTTTGA
CTACATTGGCAGCTCGAGGATGGTCTACGATATGGAGAAGGGCAAGTTTCCCGTGCAGTTAGA
GAATGTTGACTCATTTGTGGAGCTGGGACAGGTGGCCTTAAGAACTTCATTAGAGCTTTGGAT
GCACACAGATCCTGTTTCTCAGAAAAATGAGTCTGTACGGAACCAGGTGGAGGATCTCCTGGC
CACATTGGAGAAGAGTGGTGCTGGTGTCCCTGCTGTCATCCTCAGGAGGCCAAATCAGTCCCA
GCCTCTCCCACCATCTTCCCTGCAGCGATTTCTTCGAGCTCGAAACATCTCTGGCGTTGTTCT
GGCTGACCACTCTGGTGCCTTCCATAACAAATATTACCAGAGTATTTACGACACTGCTGAGAA
CATTAATGTGAGCTATCCCGAATGGCTGAGCCCTGAAGAGGACCTGAACTTTGTAACAGACAC
TGCCAAGGCCCTGGCAGATGTGGCCACGGTGCTGGGACGTGCTCTGTATGAGCTTGCAGGAGG
AACCAACTTCAGCGACACAGTTCAGGCTGATCCCCAAACGGTTACCCGCCTGCTCTATGGGTT
CCTGATTAAAGCCAACAACTCATGGTTCCAGTCTATCCTCAGGCAGGACCTAAGGTCCTACTT
GGGTGACGGGCCTCTTCAACATTACATCGCTGTCTCCAGCCCCACCAACACCACTTATGTTGT
ACAGTATGCCTTGGCAAATTTGACTGGCACAGTGGTCAACCTCACCCGAGAGCAGTGCCAGGA
TCCAAGTAAAGTCCCAAGTGAAAACAAGGATCTGTATGAGTACTCATGGGTCCAGGGCCCTTT
GCATTCTAATGAGACGGACCGACTCCCCCGGTGTGTGCGTTCTACTGCACGATTAGCCAGGGC
CTTGTCTCCTGCCTTTGAACTGAGTCAGTGGAGCTCTACTGAATACTCTACATGGACTGAGAG
CCGCTGGAAAGATATCCGTGCCCGGATATTCTCATCGCCAGCAAAGAGCTTGAGTTGATCAC
CCTGACAGTGGGCTTCGGCATCCTCATCTTCTCCCTCATCGTCACCTACTGCATCAATGCCAA
AGCTGATGTCCTTTTCATTGCTCCCGGGAGCCAGGAGCTGTGTCATAC**TGA**GGAGGACCCCA
GCTTTTCTTGCCAGNTCAGCAGTTCACTTCCTAGAGCATCTGTCCCACTGGGACACAACCACT
AATTTGTCACTGGAACCTCCCTGGGCCTGTCTCAGATTGGGATTAACATAAAAGAGTGGAACT
ATCCAAAAGAGACAGGGAGAAATAAATAAATTGCCTCCCTTCCTCCGCTCCCCTTTCCCATCA
CCCCTTCCCCATTTCCTCTTCCTTCTCTACTCATGCCAGATTTTGGGATTACAAATAGAAGCT
TCTTGCTCCTGTTTAACTCCCTAGTTACCCACCCTAATTTGCCCTTCAGGACCCTTCTACTTT
TTCCTTCCTGCCCTGTACCTCTCTCTGCTCCTCACCCCCACCCCTGTACCCAGCCACCTTCCT
GACTGGGAAGGACATAAAAGGTTTAATGTCAGGGTCAAACTACATTGAGCCCCTGAGGACAGG
GGCATCTCTGGGCTGAGCCTACTGTCTCCTTCCCACTGTCCTTTCTCCAGGCCCTCAGATGGC
ACATTAGGGTGGGCGTGCTGCGGGTGGGTATCCCACCTCCAGCCCACAGTGCTCAGTTGTACT
TTTTATTAAGCTGTAATATCTATTTTTGTTTTTGTCTTTTTCCTTTATTCTTTTTGTAAATAT
ATATATAATGAGTTTCATTAAAATAGATTATCCC

FIGURE 202

MATAGGGSGADPGSRGLLRLLSFCVLLAGLCRGNSVERKIYIPLNKTAPCVRLLNATHQIGCQ

SSISGDTGVIHVVEKEEDLQWVLTDGPNPPYMVLLESKHFTRDLMEKLKGRTSRIAGLAVSLT

KPSPASGFSPSVQCPNDGFGVYSNSYGPEFAHCREIQWNSLGNGLAYEDFSFPIFLLEDENET

KVIKQCYQDHNLSQNGSAPTFPLCAMQLFSHMHAVISTATCMRRSSIQSTFSINPEIVCDPLS

DYNVWSMLKPINTTGTLKPDDRVVVAATRLDSRSFFWNVAPGAESAVASFVTQLAAAEALQKA

PDVTTLPRNVMFVFFQGETFDYIGSSRMVYDMEKGKFPVQLENVDSFVELGQVALRTSLELWM

HTDPVSQKNESVRNQVEDLLATLEKSGAGVPAVILRRPNQSQPLPPSSLQRFLRARNISGVVL

ADHSGAFHNKYYQSIYDTAENINVSYPEWLSPEEDLNFVTDTAKALADVATVLGRALYELAGG

TNFSDTVQADPQTVTRLLYGFLIKANNSWFQSILRQDLRSYLGDGPLQHYIAVSSPTNTTYVV

QYALANLTGTVVNLTREQCQDPSKVPSENKDLYEYSWVQGPLHSNETDRLPRCVRSTARLARA

LSPAFELSQWSSTEYSTWTESRWKDIRARIFLIASKELELITLTVGFGILIESLIVTYCINAK

ADVLFIAPREPGAVSY

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 671-692

N-glycosylation sites.

amino acids 45-49, 55-59, 187-191, 200-204, 204-208, 264-268, 387-391, 417-421, 435-439, 464-468, 506-510, 530-534, 562-566, 573-577, 580-584, 612-616

Glycosaminoglycan attachment site.

amino acids 404-408

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 232-236

N-myristoylation site.

amino acids 5-11, 6-12, 9-15, 29-35, 61-67, 120-126, 146-152, 168-174, 205-211, 294-300, 438-444, 446-452, 504-510, 576-582

FIGURE 203

```
GCTAGACCGAGCCCTGGGAGGCTACGGGCTCCCCCGGAAACCCTGCCAGGGGAGCCGGGTTTT
GAGCTCAGGCGCCTCTAGCGGCGGCCCCCAGAAATCTGACTCGCGAGGCCAGAGTTGCAGGGA
CTGAATAGCAAACTGAGGCTGAGTAGGGAACAGACCATGAGGTCAGTGCAGATCTTCCTCTCC
CAATGCCGTTTGCTCCTTCTACTAGTTCCGACAATGCTCCTTAAGTCTCTTGGCGAAGATGTA
ATTTTTCACCCTGAAGGGGAGTTTGACTCGTATGAAGTCACCATTCCTGAGAAGCTGAGCTTC
CGGGGAGAGGTGCAGGGTGTGGTCAGTCCCGTGTCCTACCTACTGCAGTTAAAAGGCAAGAAG
CACGTCCTCCATTTGTGGCCCAAGAGACTTCTGTTGCCCCGACATCTGCGCGTTTTCTCCTTC
ACAGAACATGGGGAACTGCTGGAGGATCATCCTTACATACCAAAGGACTGCAACTACATGGGC
TCCGTGAAAGAGTCTCTGGACTCTAAAGCTACTATAAGCACATGCATGGGGGGTCTCCGAGGT
GTATTTAACATTGATGCCAAACATTACCAAATTGAGCCCCTCAAGGCCTCTCCCAGTTTTGAA
CATGTCGTCTATCTCCTGAAGAAAGAGCAGTTTGGGAATCAGGTTTGTGGCTTAAGTGATGAT
GAAATAGAATGGCAGATGGCCCCTTATGAGAATAAGGCGAGGCTAAGGGACTTTCCTGGATCC
TATAAACACCCAAAGTACTTGGAATTGATCCTACTCTTTGATCAAAGTAGGTATAGGTTTGTG
AACAACAATCTTTCTCAAGTCATACATGATGCCATTCTTTTGACTGGGATTATGGACACCTAC
TTTCAAGATGTTCGTATGAGGATACACTTAAAGGCTCTTGAAGTATGGACAGATTTTAACAAA
ATACGCGTTGGATATCCAGAGTTAGCTGAAGTTTTAGGCAGATTTGTAATATATAAAAAAAGT
GTATTAAATGCTCGCCTGTCATCAGATTGGGCACATTTATATCTTCAAAGAAAATATAATGAT
GCTCTTGCATGGTCGTTTGGAAAAGTGTGTTCTCTAGAATATGCTGGATCAGTGAGTACTTTA
CTAGATACAAATATCCTTGCCCCTGCTACCTGGTCTGCTCATGAGCTGGGTCATGCTGTAGGA
ATGTCACATGATGAACAATACTGCCAATGTAGGGGTAGGCTTAATTGCATCATGGGCTCAGGA
CGCACTGGGTTTAGCAATTGCAGTTATATCTCTTTTTTTAAACATATCTCTTCGGGAGCAACA
TGTCTAAATAATATCCCAGGACTAGGTTATGTGCTTAAGAGATGTGGAAACAAAATTGTGGAG
GACAATGAGGAATGTGACTGTGGTTCCACAGAGGAGTGTCAGAAAGATCGGTGTTGCCAATCA
AATTGTAAGTTGCAACCAGGTGCCAACTGTAGCATTGGACTTTGCTGTCATGATTGTCGGTTT
CGTCCATCTGGATACGTGTGTAGGCAGGAAGGAAATGAATGTGACCTTGCAGAGTACTGCGAC
GGGAATTCAAGTTCCTGCCCAAATGACGTTTATAAGCAGGATGGAACCCCTTGCAAGTATGAA
GGCCGTTGTTTCAGGAAGGGGTGCAGATCCAGATATATGCAGTGCCAAAGCATTTTTGGACCT
GATGCCATGGAGGCTCCTAGTGAGTGCTATGATGCAGTTAACTTAATAGGTGATCAATTTGGT
AACTGTGAGATTACAGGAATTCGAAATTTTAAAAAGTGTGAAAGTGCAAATTCAATATGTGGC
AGGCTACAGTGTATAAATGTTGAAACCATCCCTGATTTGCCAGAGCATACGACTATAATTTCT
ACTCATTTACAGGCAGAAAATCTCATGTGCTGGGGCACAGGCTATCATCTATCCATGAAACCC
ATGGGAATACCTGACCTAGGTATGATAAATGATGGCACCTCCTGTGGAGAAGGCCGGGTATGT
TTTAAAAAAATTGCGTCAATAGCTCAGTCCTGCAGTTTGACTGTTTGCCTGAGAAATGCAAT
ACCCGGGGTGTTTGCAACAACAGAAAAAACTGCCACTGCATGTATGGGTGGGCACCTCCATTC
TGTGAGGAAGTGGGGTATGGAGGAAGCATTGACAGTGGGCCTCCAGGACTGCTCAGAGGGGCG
ATTCCCTCGTCAATTTGGGTTGTGTCCATCATAATGTTTCGCCTTATTTTATTAATCCTTTCA
GTGGTTTTTGTGTTTTTCCGGCAAGTGATAGGAAACCACTTAAAACCCAAACAGGAAAAAATG
CCACTATCCAAAGCAAAAACTGAACAGGAAGAATCTAAAACAAAAACTGTACAGGAAGAATCT
AAAACAAAAACTGGACAGGAAGAATCTGAAGCAAAAACTGGACAGGAAGAATCTAAAGCAAAA
ACTGGACAGGAAGAATCTAAAGCAAACATTGAAAGTAAACGACCCAAAGCAAGAGTGTCAAG
AAACAAAAAAGTAACCGGGCAATCCATACTCATTCAGTAACACAGGCTCATTTATTTAACCA
GCTAATCATTTATCCAAAGGCTTTCCATTCTTCTCCCAATATTTTTTTACTTTAATTTTTCCC
ACAAGTTTTGATCAGCAAATAAACAGCATTCTTGTTTTGGAAACAAAAA
```

FIGURE 204

MRSVQIFLSQCRLLLLLVPTMLLKSLGEDVIFHPEGEFDSYEVTIPEKLSFRGEVQGVVSPVS

YLLQLKGKKHVLHLWPKRLLLPRHLRVFSFTEHGELLEDHPYIPKDCNYMGSVKESLDSKATI

STCMGGLRGVFNIDAKHYQIEPLKASPSFEHVVYLLKKEQFGNQVCGLSDDEIEWQMAPYENK

ARLRDFPGSYKHPKYLELILLFDQSRYRFVNNNLSQVIHDAILLTGIMDTYFQDVRMRIHLKA

LEVWTDFNKIRVGYPELAEVLGRFVIYKKSVLNARLSSDWAHLYLQRKYNDALAWSFGKVCSL

EYAGSVSTLLDTNILAPATWSAHELGHAVGMSHDEQYCQCRGRLNCIMGSGRTGFSNCSYISF

FKHISSGATCLNNIPGLGYVLKRCGNKIVEDNEECDCGSTEECQKDRCCQSNCKLQPGANCSI

GLCCHDCRFRPSGYVCRQEGNECDLAEYCDGNSSSCPNDVYKQDGTPCKYEGRCFRKGCRSRY

MQCQSIFGPDAMEAPSECYDAVNLIGDQFGNCEITGIRNFKKCESANSICGRLQCINVETIPD

LPEHTTIISTHLQAENLMCWGTGYHLSMKPMGIPDLGMINDGTSCGEGRVCFKKNCVNSSVLQ

FDCLPEKCNTRGVCNNRKNCHCMYGWAPPFCEEVGYGGSIDSGPPGLLRGAIPSSIWVVSIIM

FRLILLILSVVFVFFRQVIGNHLKPKQEKMPLSKAKTEQEESKTKTVQEESKTKTGQEESEAK

TGQEESKAKTGQEESKANIESKRPKAKSVKKQKK

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 684-705

N-glycosylation sites.

amino acids 222-226, 372-376, 438-442, 473-477, 625-629

N-myristoylation sites.

amino acids 131-137, 168-174, 235-241, 319-325, 364-370, 436-442, 472-478, 609-615, 642-648, 668-674, 676-680, 680-686, 749-755, 758-764, 767-773

Amidation site.

amino acids 69-73

Disintegrins proteins

amino acids 429-479

EGF-like domain proteins

amino acids 650-662

Neutral zinc metallopeptidases, zinc-binding region proteins

amino acids 335-345

FIGURE 205

```
CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGGAAGGTTGAATGGGGTAGAAGGCCTG
TTGTGGAGGGAAACCACCCATCCTCCTGCCTCCCACCACCACCATCATCCTGGCTGGACGGAG
AGGGTGACGGGGGCTGGGAAGGGGCAGCTCATGTTCAGGTTTCCAGGAGGGGCTACCTGTTGA
CTGTCTTTGCAGGAAGAAGAAACACCTGAGTGACCAGATGTCCCAGCTCCAGGTGCCTTGCC
AGATGGCCAGAACCACACCTCTTGAAGAGTGACAGTGCTGTGGAGCATGGTTTCTGCACACCT
GGAATGACTGGAACCCCAAAGACTCAAGAAGGAGCTAAAGATCTTGAAGTAGACATGAATAAA
ACAGAAGGCTGTGGACCACCTGTCGAGATGGAGAAGTCCTTCTGAGGCTATCCAAACACGGAC
CAGGCCATGAGACCCCG**ATG**ACCATCCCTGAATTTTTTCGAGAGTCAGTCAACCGATTTGGAA
CTTATCCAGCCCTCCCATCCAAGAATGGCAAAAAGTGGGAAATTCTGAATTTCAACCAGTACT
ATGAGGCTTGTCGGAAGGCTGCAAAATCCTTGATCAAGCTGGGTTTGGAGCGTTTCCACGGAG
TTGGTATCCTGGGGTTTAACTCTGCAGAGTGGTTTATCACTGCTGTTGGTGCCATCCTAGCCG
GGGGTCTTTGTGTTGGTATTTATGCCACCAACTCTGCCGAGGCTTGTCAATATGTCATCACTC
ATGCCAAAGTGAACATCTTGCTGGTTGAGAATGATCAACAGTTACAGAAAATCCTTTCGATTC
CACAGAGCAGCCTAGAGCCCCTAAAAGCGATCATCCAGTACAGACTGCCAATGAAGAAGAACA
ACAACTTGTACTCTTGGGATGATTTCATGGAACTTGGCAGAAGTATCCCTGACACCCAACTGG
AGCAGGTCATCGAGAGCCAGAAGGCGAATCAATGCGCAGTGCTCATCTACACTTCAGGGACCA
CAGGCATACCCAAGGGAGTGATGCTCAGTCATGACAACATCACGTGGATTGCAGGAGCAGTGA
CAAAGGACTTTAAACTGACAGACAAGCATGAGACGGTGGTTAGCTACCTCCCACTCAGCCATA
TTGCAGCACAGATGATGGACATCTGGGTACCCATAAAGATTGGGGCGCTCACATACTTTGCTC
AAGCAGATGCTCTCAAGGGCACCTTGGTAAGTACTCTAAAGGAGGTAAAACCTACTGTCTTCA
TTGGAGTGCCTCAAATTTGGGAGAAGATACATGAGATGGTGAAGAAAAATAGTGCCAAGTCCA
TGGGCTTGAAGAAGAAGGCATTCGTGTGGGCAAGAAACATTGGCTTCAAGGTCAACTCAAAAA
AGATGTTGGGGAAATATAATACTCCCGTGAGCTACCGCATGGCTAAGACTCTCGTGTTCAGCA
AAGTCAAGACATCCCTTGGCTTGGATCACTGTCACTCTTTTATCAGTGGGACTGCGCCCCTCA
ACCAAGAGACTGCCGAGTTCTTTCTAAGCTTGGACATACCTATAGGCGAGTTGTATGGGTTGA
GTGAGAGCTCGGGACCCCACACGATATCCAACCAGAATAACTACAGGCTTCTAAGCTGTGGCA
AGATCTTGACTGGGTGTAAGAATATGCTGTTCCAGCAGAACAAGGATGGCATTGGGGAGATCT
GCCTCTGGGGTAGGCACATCTTCATGGGCTATCTGGAAAGTGAGACTGAAACTACAGAGGCCA
TCGATGATGAAGGCTGGCTACACTCTGGGGATCTGGGCCAGCTGGACGGTCTGGGTTTCCTCT
ATGTCACCGGCCACATCAAAGAAATCCTTATCACTGCTGGTGGTGAAAATGTGCCCCCCATTC
CTGTTGAGACCTTGGTTAAGAAGAAGATCCCCATCATCAGTAACGCCATGTTAGTAGGAGATA
AACTGAAGTTTCTGAGCATGTTGCTGACGCTGAAGTGTGAGATGAATCAGATGAGCGGAGAAC
CTCTGGACAAGCTGAACTTCGAGGCCATCAACTTCTGTCGGGGTCTGGGCAGCCAGGCATCCA
CCGTGACTGAGATTGTGAAGCAGCAAGACCCCCTGGTCTACAAGGCCATCCAGCAAGGCATCA
ATGCTGTGAACCAGGAAGCCATGAACAATGCACAGAGGATTGAAAAGTGGGTCATCTTGGAGA
AGGACTTTTCCATCTATGGTGGAGAGCTAGGTCCAATGATGAAACTTAAGAGACATTTTGTAG
CCCAGAAATACAAAAAACAAATTGATCACATGTACCAC**TGA**CTGCTTTGATGGAGCTGCTCTC
AGCTGTTCTGATGCCTTCAGCAGGAAGACCTCATTGCAATAAGTGAAATGCTGCTCTAGGTAG
AAGCTCTCCCTGCTGTTTTTAAGAAGCCACATTCCTCATTGGTCAGTTTCTTGATTGTTCGTC
TGTTGGAGAGGTGCTCCCTAGAAGAACCTGCCATACGTTTCAAAGCAATAAATCACTGTATA
TCTTTCTAAGGACCTTCAAGTCATGACTCCAGGGAAGCCTATTGGGAAGTCTACTAAAAACTG
CCTGATTTACAAGAAAGACCTGAACTTGTGGGCTCCCATTTGATTTTTTTCTCCTCAGGGGAC
TCAGACATTAGAAAGAAAAGCCTCACAGATTTGAAGAACTGGACCCCCAAATCAACTCACCT
GCCTGGAAGCAACTGGGAAACCCTTCCAATAAGTCCTGATAATAAAGCACTTCAGGGTCCCAA
AAAAAAAAAA
```

FIGURE 206

MTIPEFFRESVNRFGTYPALPSKNGKKWEILNFNQYYEACRKAAKSLIKLGLERFHGVGILGF

NSAEWFITAVGAILAGGLCVGIYATNSAEACQYVITHAKVNILLVENDQQLQKILSIPQSSLE

PLKAIIQYRLPMKKNNNLYSWDDFMELGRSIPDTQLEQVIESQKANQCAVLIYTSGTTGIPKG

VMLSHDNITWIAGAVTKDFKLTDKHETVVSYLPLSHIAAQMMDIWVPIKIGALTYFAQADALK

GTLVSTLKEVKPTVFIGVPQIWEKIHEMVKKNSAKSMGLKKKAFVWARNIGFKVNSKKMLGKY

NTPVSYRMAKTLVFSKVKTSLGLDHCHSFISGTAPLNQETAEFFLSLDIPIGELYGLSESSGP

HTISNQNNYRLLSCGKILTGCKNMLFQQNKDGIGEICLWGRHIFMGYLESETETTEAIDDEGW

LHSGDLGQLDGLGFLYVTGHIKEILITAGGENVPPIPVETLVKKKIPIISNAMLVGDKLKFLS

MLLTLKCEMNQMSGEPLDKLNFEAINFCRGLGSQASTVTEIVKQQDPLVYKAIQQGINAVNQE

AMNNAQRIEKWVILEKDFSIYGGELGPMMKLKRHFVAQKYKKQIDHMYH

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 65-86

N-glycosylation site.

amino acids 196-200

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 282-286

Tyrosine kinase phosphorylation sites.

amino acids 547-555, 608-616

N-myristoylation sites.

amino acids 15-21, 74-80, 80-86, 84-90, 185-191, 189-195, 253-259, 337-343, 371-377, 448-454, 536-542

Amidation site.

amino acids 24-28

Putative AMP-binding domain signature.

amino acids 177-189

Putative AMP-binding domain proteins.

amino acids 173-190

FIGURE 207

CCCACGCGTCCGCCCACGCGTCCGCGGACGCGTGGGGCCAGATCGCGGCCGGCGCCAGCGCCA
CCGTCCGGTCCACCCGCCAGCCCGCACAGCCGCGCCGCCGCCGAGCGTTTCGTGAGCGGCGCT
CCGAGGATCAGGAATGGGGCTTCGGGCGCTGGGCGCGCTCCGAACCCGGCGCACGTAAGAGCC
TGGGAGCGCCCGAGCCGCCCGGCTGCCCGGAGCCCCATCGCCTAGGACCGGGAGATGCTGGAA
ATGCAACCGCCTGTTCCCCGAGGAGCCGCTGCCCCCGGGACCCCCTGGCACTGTGCGCACCCT
GGTCAGCAGCCCCGGAGAAGACGGCGCCCCCAACGCCCGACCCGCGTGGCCGTGGCAGCGCC
ACGCGAGCCCTCTAGGCGACCGCAGGGCCACAGCAGCTCAGCCGCCGGTGCCCCCTCGGAAAC
CATGACCCCCGGCGCGGGCCCATGGAGCC**ATG**GCCTATAGGGTCCTGGGCCGCGCGGGGCCAC
CTCAGCCGCGGAGGGCGCGCAGGCTGCTCTTCGCCTTCACGCTCTCGCTCTCCTGCACTTACC
TGTGTTACAGCTTCCTGTGCTGCTGCGACGACCTGGGTCGGAGCCGCCTCCTCGGCGCGCCTC
GCTGCCTCCGCGGCCCCAGCGCGGGCGGCCAGAAACTTCTCCAGAAGTCCCGCCCCTGTGATC
CCTCCGGGCCGACGCCCAGCGAGCCCAGCGCTCCCAGCGCGCCCGCCGCCGCCGTGCCCGCCC
CTCGCCTCTCCGGTTCCAACCACTCCGGCTCACCCAAGCTGGGTACCAAGCGGTTGCCCCAAG
CCCTCATTGTGGGCGTGAAGAAGGGGGGCACCCGGGCCGTGCTGGAGTTTATCCGAGTACACC
CGGACGTGCGGGCCTTGGGCACGGAACCCCACTTCTTTGACAGGAACTACGGCCGCGGGCTGG
ATTGGTACAGGAGCCTGATGCCCAGGACCCTCGAGAGCCAGATCACGCTGGAGAAGACGCCCA
GCTACTTTGTCACTCAAGAGGCTCCTCGACGCATCTTCAACATGTCCCGAGACACCAAGCTGA
TCGTGGTTGTGCGGAACCCTGTGACCCGTGCCATCTCTGATTACACGCAGACACTCTCCAAGA
AGCCCGACATCCCGACCTTTGAGGGCCTCTCCTTCCGCAACCGCACCCTGGGCCTGGTGGACG
TGTCATGGAACGCCATCCGCATCGGCATGTACGTGCTGCACCTGGAGAGCTGGCTGCAGTACT
TCCCGCTAGCTCAGATTCACTTCGTCAGTGGCGAGCGACTCATCACTGACCCGGCCGGCGAGA
TGGGGCGAGTCCAGGACTTCCTGGGCATTAAGAGATTCATCACGGACAAGCACTTCTATTTCA
ACAAGACCAAAGGATTCCCTTGCTTGAAAAAAACAGAATCGAGCCTCCTGCCTCGATGCTTGG
GCAAATCAAAAGGGAGAACTCATGTACAGATTGATCCTGAAGTGATAGACCAGCTCCGAGAAT
TTTATAGACCGTATAATATCAAATTTTATGAAACCGTTGGGCAGGACTTCAGGTGGGAA**TAA**G
CCCACGAAAGGAAAGGGCTCTCAAGGGCTCTTCTGCTCATCTCTTCCGTGAGATTTGCTCCCA
GACCCTCTGATCTCCCTCCAACAAACCCTGGCTCCAGCCCCCTTTCCCAACTTGAGTTGCATC
ATCTTGGAACCAGGAAGCCCAGCTAAAGCCAAGAGACCAGAGAGTCCCTGCCACTAGTTTTCA
TCAGTCTGTTCAAGCAAAGTTGATCTGCTCCTGGCACGTCCAGTAAATTCCAGAATCATTCTC
CTTTCTGCCCATAAAGGGCCTTGGAGAATTGCTTTAAGAAGAGTGAATGTTCCAATGATGATA
GATATTATAAGCGATGATGGTTCTGTTGCTATGAACACAGCAGTCGGTCCCTGTCATTGTCCA
CCCAGGAGTGGCCTTGTTAATTCCAAGTGGCATGTATCTTCCCTCTGAGCTTCATTTCTTCAA
GATGCTCTGGGTGGTGGGATGGGAGACCATCCTCAGCCCTCCTCAGACCTTATCAATTCATTG
AGAGATTGCAAAGCTGAAAGCACCTCCGGCCACTCCTGGGAGACAGACCCTTTGGTGATGAAA
TAAACCAGTGACTTCAGAGCCTATGGTCTCAACTGTGCTTGAAAAACACTGTCTCTGAAAACA
ACTTTGTGATTCTCCCTGCTCCCTGTGGACAAAAGCACATAATTCTGCTGTTACGGGTACTTT
GCTCATACGAGCTTTCATGTTCAGCATGCAATGGAATCATGCTTGTCCATGTGAAATAAATAT
GGCTCTCTCGTGTCCTTAATGCTGGGCTTTTCTCTGTAAGCTGGTTCTGCAGCACAATTCATT
AATTAAACTTCTCCCAGTGCAAGAAGGCAGCTGGTGCTGGGGGTGGTCTGGGGGGTCAGGGAG
GAGGGCAAGGACTACATGGGGCAGAGGCAAGGCGGTGGTGGAGATGAGGAAGAAGTTCTTCT
TGGCAGAAGCTGGGGCAGAAAGATCACATGAGATCTGTGGGGACACCCTCTATCTGAAACATA
AGTCTGTGTTCATTCTCTGCTTAGAAATTTTAGATCTGAAGTGCTACACTGAAGGTCCGAAGG
TTGATGGGGCATCAGATATCTTTTTGGTTGGCCAGCATGATATTTTGAAATAACTGTCAACAG
TTAGAAACTGGGAGCATTCATATGTAAAAAATATGGATTTTCAGCTTCTTCTTAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAA

FIGURE 208

MAYRVLGRAGPPQPRRARRLLFAFTLSLSCTYLCYSFLCCCDDLGRSRLLGAPRCLRGPSAGG
QKLLQKSRPCDPSGPTPSEPSAPSAPAAAVPAPRLSGSNHSGSPKLGTKRLPQALIVGVKKGG
TRAVLEFIRVHPDVRALGTEPHFFDRNYGRGLDWYRSLMPRTLESQITLEKTPSYFVTQEAPR
RIFNMSRDTKLIVVVRNPVTRAISDYTQTLSKKPDIPTFEGLSFRNRTLGLVDVSWNAIRIGM
YVLHLESWLQYFPLAQIHFVSGERLITDPAGEMGRVQDFLGIKRFITDKHFYFNKTKGFPCLK
KTESSLLPRCLGKSKGRTHVQIDPEVIDQLREFYRPYNIKFYETVGQDFRWE

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 102-106, 193-197, 235-239, 306-310

Tyrosine kinase phosphorylation site.

amino acids 296-305

N-myristoylation sites.

amino acids 51-57, 100-106, 121-127, 125-131

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 20-31

FIGURE 209

```
CTTTCCTTATCTGTGTGTACTCTTATCTCACTGTTCTATTTTTTCTCCTCATTTATATTAACT
CTTTCTTACCTTTTTTTTCTGAACTTCTAGGCCTTCTCTTTCCAGAACTGGTGGAAGACAAATG
AAACGGCCAAGATGGTAAGAAACAAGCCGCATTTCTCCTTGGGGAGACTGATAATTTAAAAGG
TTTGTTGTGTCAGAAACATTCCCAGCTTCATCACCAACCCTTTCCTTCCACCTCTGCCCACTG
GAGACCACTTACATCCCGAAGCGGACGCGGCAGCTGAAGTCAGGAAACCATGCATCACATTAG
CAGGAGCCAACTGCAGACTTTAAACTCCGTTCAACATGTGGATGCGGCAGAGAA**ATG**ACCTGT
CCAGACAAGCCGGGGCAGCTCATAAACTGGTTCATCTGCTCCCTGTGCGTCCCGCGGGTGCGT
AAGCTCTGGAGCAGCCGGCGTCCAAGGACCCGGAGAAACCTTCTGCTGGGCACTGCGTGTGCC
ATCTACTTGGGCTTCCTGGTGAGCCAGGTGGGGAGGGCCTCTCTCCAGCATGGACAGGCGGCT
GAGAAGGGGCCACATCGCAGCCGCGACACCGCCGAGCCATCCTTCCCTGAGATACCCCTGGAT
GGTACCCTGGCCCCTCCAGAGTCCCAGGGCAATGGGTCCACTCTGCAGCCCAATGTGGTGTAC
ATTACCCTACGCTCCAAGCGCAGCAAGCCGGCCAATATCCGTGGCACCGTGAAGCCCAAGCGC
AGGAAAAAGCATGCAGTGGCATCGGCTGCCCCAGGGCAGGAGGCTTTGGTCGGACCATCCCTT
CAGCCGCAGGAAGCGGCAAGGGAAGCTGATGCTGTAGCACCTGGGTACGCTCAGGGAGCAAAC
CTGGTTAAGATTGGAGAGCGACCCTGGAGGTTGGTGCGGGGTCCGGGAGTGCGAGCCGGGGGC
CCAGACTTCCTGCAGCCCAGCTCCAGGGAGAGCAACATTAGGATCTACAGCGAGAGCGCCCCC
TCCTGGCTGAGCAAAGATGACATCCGAAGAATGCGACTCTTGGCGGACAGCGCAGTGGCAGGG
CTCCGGCCTGTGTCCTCTAGGAGCGGAGCCCGTTTGCTGGTGCTGGAGGGGGGCGCACCTGGC
GCTGTGCTCCGCTGTGGCCCTAGCCCCTGTGGGCTTCTCAAGCAGCCCTTGGACATGAGTGAG
GTGTTTGCCTTCCACCTAGACAGGATCCTGGGGCTCAACAGGACCCTGCCGTCTGTGAGCAGG
AAAGCAGAGTTCATCCAAGATGGCCGCCCATGCCCCATCATTCTTTGGGATGCATCTTTATCT
TCAGCAAGTAATGACACCCATTCTTCTGTTAAGCTCACCTGGGGAACTTATCAGCAGTTGCTG
AAACAGAAATGCTGGCAGAATGGCCGAGTACCCAAGCCTGAATCAGGTTGTACTGAAATACAT
CATCATGAGTGGTCCAAGATGGCACTCTTTGATTTTTTGTTACAGATTTATAATCGCTTAGAT
ACAAATTGCTGTGGATTCAGACCTCGCAAGGAAGATGCCTGTGTACAGAATGGATTGAGGCCA
AAATGTGATGACCAAGGTTCTGCGGCTCTAGCACACATTATCCAGCGAAAGCATGACCCAAGG
CATTTGGTTTTTATAGACAACAAGGGTTTCTTTGACAGGAGTGAAGATAACTTAAACTTCAAA
TTGTTAGAAGGCATCAAAGAGTTTCCAGCTTCTGCAGTTTCTGTTTTGAAGAGCCAGCACTTA
CGGCAGAAACTTCTTCAGTCTCTGTTTCTTGATAAAGTGTATTGGGAAAGTCAAGGAGGTAGA
CAAGGAATTGAAAGCTTATCGATGTAATAGAACACAGAGCCAAAATTCTTATCACCTATATC
AATGCACACGGGGTCAAAGTATTACCTATGAATGAA**TGA**CAAAAGAATCTTCTGGCTAGGGTG
TTAGATATATTTATGCATTTTTGGTTTTGTTTTAAATCAAGCACATCAACCTCAAGCCCGTT
TAGCAATGAGGCAGTGTAGATGAATACGTAAAATAAATGACTTTAACCAAGTAGCTATAAAGG
GACTTAGCACTGTATGCATACTTAAAAAGGTTTTGAAAAACAAACTACTTGAGAAATATTTGT
TTATATTTTTCTCAACATCATGCTATGTGTCAGTCTGAACATCTGACAACAGAAATTTCAGT
TATTATTCTAGCTAAGTTTTGAAAACATTTGTCATGCTGTTTAATAGAAACTGCAAACCAGA
GATACTGACTCCATTAATAAACCATATTTTGTGCCGTTTTGACTGTTCTGACCAAATACTAAT
GGGAACAATTCTTGACGTTTTTCTGTTGCTGATTGTTAACATAGAGCAGTCTCTACACTACCC
TGAGGCAACTCTACATTGGAACACTGAGGCTTACAGCCTGCAAGAGCATCAGAGCTGACCATA
CATTTAAACAGAAATGCTGGTTTATTTGCAAATCACCAGTATATTTTCTATTGTGTCTATAA
AAAATCAGTCATTTAAGTACAAGAATCATATTTTCCATTCCTTTTTAGAAATTTATTTTGTTG
TCCCTATGGAAATCATTCACATCTGACAATTTATATGTTAAAGAGTTTTACTCTCTCTATTTT
GGTCCAATTTGTATCTAGTGGCTGAGAAATTAAATAATTCTAAAGTATGAAGTTACCTATCTG
AAAATGTACTTACAGAGTATCATTTTAAAATGGATGTCTCTTTAAAAATTTTGTTACTTTTAC
CAACAATGTAATATAATTTATGTATATTTTATTAATAATAGTGAATTCCTTAAAATTTGTTCT
ATGTACTTATATTTAATTTGATTTAATGGTTACTGCCCAGATATTGAGAAATGGTTCAAATAT
TGAGTGTGTTTCAATAA
```

FIGURE 210

MTCPDKPGQLINWFICSLCVPRVRKLWSSRRPRTRRNLLLGTACAIYLGFLVSQVGRASLQHG

QAAEKGPHRSRDTAEPSFPEIPLDGTLAPPESQGNGSTLQPNVVYITLRSKRSKPANIRGTVK

PKRRKKHAVASAAPGQEALVGPSLQPQEAAREADAVAPGYAQGANLVKIGERPWRLVRGPGVR

AGGPDFLQPSSRESNIRIYSESAPSWLSKDDIRRMRLLADSAVAGLRPVSSRSGARLLVLEGG

APGAVLRCGPSPCGLLKQPLDMSEVFAFHLDRILGLNRTLPSVSRKAEFIQDGRPCPIILWDA

SLSSASNDTHSSVKLTWGTYQQLLKQKCWQNGRVPKPESGCTEIHHHEWSKMALFDFLLQIYN

RLDTNCCGFRPRKEDACVQNGLRPKCDDQGSAALAHIIQRKHDPRHLVFIDNKGFFDRSEDNL

NFKLLEGIKEFPASAVSVLKSQHLRQKLLQSLFLDKVYWESQGGRQGIEKLIDVIEHRAKILI

TYINAHGVKVLPMNE

Transmembrane domain:

amino acids 40-56

N-glycosylation sites.

amino acids 98-102, 289-293, 322-326

N-myristoylation sites.

amino acids 8-14, 41-47, 97-103, 187-193, 251-257, 252-258, 287-293, 484-490

FIGURE 211

```
GTGGGGTGGTGAGCGCAGCGCCGAGGATGAGGAGGTGCAACAGCGGCTCCGGGCCGCCGCCGTCGCTGCTGCTGC
TGCTGCTGTGGCTGCTCGCGGTTCCCGGCGCTAACGCGGCCCCGCGGTCGGCGCTCTATTCGCCTTCCGACCCGC
TGACGCTGCTGCAGGCGGACACGGTGCGCGGCGCGGTGCTGGGCTCCCGCAGCGCCTGGGCCGTGGAGTTCTTCG
CCTCCTGGTGCGGCCACTGCATCGCCTTCGCCCCGACGTGGAAGGCGCTGGCCGAAGACGTCAAAGCCTGGAGGC
CGGCCCTGTATCTCGCCGCCCTGGACTGTGCTGAGGAGACCAACAGTGCAGTCTGCAGAGACTTCAACATCCCTG
GCTTCCCGACTGTGAGGTTCTTCAAGGCCTTTACCAAGAACGGCTCGGGAGCAGTATTTCCAGTGGCTGGTGCTG
ACGTGCAGACGCTGCGGGAGAGGCTCATTGACGCCCTGGAGTCCCATCATGACACGTGGCCCCCAGCCTGTCCCC
CACTGGAGCCTGCCAAGCTGGAGGAGATTGATGGATTCTTTGCGAGAAATAACGAAGAGTACCTGGCTCTGATCT
TTGAAAAGGGAGGCTCCTACCTGGGTAGAGAGGTGGCTCTGGACCTGTCCCAGCACAAAGGCGTGGCGGTGCGCA
GGGTGCTGAACACAGAGGCCAATGTGGTGAGAAAGTTTGGTGTCACCGACTTCCCCTCTTGCTACCTGCTGTTCC
GGAATGGCTCTGTCTCCCGAGTCCCCGTGCTCATGGAATCCAGGTCCTTCTATACCGCTTACCTGCAGAGACTCT
CTGGGCTCACCAGGGAGGCTGCCCAGACCACAGTTGCACCAACCACTGCTAACAAGATAGCTCCCACTGTTTGGA
AATTGGCAGATCGCTCCAAGATCTACATGGTGACCTGGAATCTGCACTGCACTACATCCTGCGGATAGAAGTGG
GCAGGTTCCCGGTCCTGGAAGGGCAGCGCCTGGTGGCCCTGAAAAAGTTTGTGGCAGTGCTGGCCAAGTATTTCC
CTGGCCGGCCCTTAGTCCAGAACTTCCTGCACTCCGTGAATGAATGGCTCAAGAGGCAGAAGAGAAATAAAATTC
CCTACAGTTTCTTTAAAACTGCCCTGGACGACAGGAAAGAGGGTGCCGTTCTTGCCAAGAAGGTGAACTGGATTG
GCTGCCAGGGGAGTGAGCCGCATTTCCGGGGCTTTCCCTGCTCCCTGTGGGTCCTCTTCCACTTCTTGACTGTGC
AGGCAGCTCGGCAAAATGTAGACCACTCACAGGAAGCAGCCAAGGCCAAGGAGGTCCTCCCAGCCATCCGAGGCT
ACGTGCACTACTTCTTCGGCTGCCGAGACTGCGCTAGCCACTTCGAGCAGATGGCTGCTGCCTCCATGCACCGGG
TGGGGAGTCCCAACGCCGCTGTCCTCTGGCTCTGGTCTAGCCACAACAGGGTCAATGCTCGCCTTGCAGGTGCCC
CCAGCGAGGACCCCCAGTTCCCCAAGGTGCAGTGGCCACCCCGTGAACTTTGTTCTGCCTGCCACAATGAACGCC
TGGATGTGCCCGTGTGGGACGTGGAAGCCACCCTCAACTTCCTCAAGGCCCACTTCTCCCCAAGCAACATCATCC
TGGACTTCCCTGCAGCTGGGTCAGCTGCCCGGAGGGATGTGCAGAATGTGGCAGCCGCCCCAGAGCTGGCGATGG
GAGCCCTGGAGCTGGAAAGCCGGAATTCAACTCTGGACCCTGGGAAGCCTGAGATGATGAAGTCCCCCACAAACA
CCACCCCACATGTGCCGGCTGAGGGACCTGAGGCAAGTCGACCCCCGAAGCTGCACCCTGGCCTCAGAGCTGCAC
CAGGCCAGGAGCCTCCTGAGCACATGGCAGAGCTTCAGAGGAATGAGCAGGAGCAGCCGCTTGGGCAGTGGCACT
TGAGCAAGCGAGACACAGGGGCTGCATTGCTGGCTGAGTCCAGGGCTGAGAAGAACCGCCTCTGGGGCCCTTTGG
AGGTCAGGCGCGTGGGCCGCAGCTCCAAGCAGCTGGTCGACATCCCTGAGGGCCAGCTGGAGGCCCGAGCTGGAC
GGGGCCGAGGCCAGTGGCTGCAGGTGCTGGGAGGGGGCTTCTCTTACCTGGACATCAGCCTCTGTGTGGGGCTCT
ATTCCCTGTCCTTCATGGGCCTGCTGGCCATGTACACCTACTTCCAGGCCAAGATAAGGGCCCTGAAGGGCCATG
CTGGCCACCCTGCAGCCTGAACCACCTGGGGAGGAGGCGGGAGAGGGAGCTGCCATCTCTAGGCACCTCAAGCCC
CCTGACCCCATTCCCTCCCCTCCCACCCCTTGCTCCTTGTCTGGCCTAGAAGTGTGGGAAATTCAGGAAAACGAG
TTGCTCCAGTGAAGCTTCTTGGGGTTGCTAGGACAGAGAGCTCCTTTGACACAAAAGACAGGAGCAGGGTCCAGG
TTCCCCTGCTGTGCAGGGAGGGCAGCCCCGGGCAGTGGGCATAGGGCAGCTCAGTCCCTGGCCTCTTAGCACCAC
ATTCCTGTTTTTTCAGCTTATTTGAAGTCCTGCCTCATTCTCACTGGAGCCTCAGTCTCTCCTGCTTGGTCTTGGC
CCTCAACTGGGGCAAGTGAAGCCAGAGGAGGGTCCCCCAGCTGGGTGGGCTGGAATGGAACTCCTCACTAGCTGC
TGGGGCTCCGCCCACCCTGCTCCCTTCCGGACAATGAAGAAGCCTTTGCACCCTGGGAGGAAGGACCACCCCGGG
CCCTCTATGCCTGGCCAGCCTCCAGCTCCTCAGACCTCCTGGGTGGGGTTTGGCTTCAGGGTGGGGTTTGGAAGC
TTCTGGAAGTCGTGCTGGTCTCCCAGGTGAGGCAAGCCATGGTTGCTGGGCTGTAGGGTGAGTGGCTTGCTTGGT
GGGACCTGACGAGTTGGTGGCATGGGAAGGATGTGGGTCTCTAGTGCCTTGCCCTGGCTTAGCTGCAGGAGAAGA
TGGCTGCTTTCACTTCCCCCCATTGAGCTCTGCTCCCTCTGAGCCTGGTCTTTTGTCCTTTTTTATTTTGGTCTC
CAAGATGAATGCTCATCTTTGGAGGGTGCCAGGTAGAAGCTAGGGAGGGGAGTGTCTTCTCTCTCCAGGTTTCAC
CTTCCAGTGTGCAGAAGTTAGAAGGGTCTGGCGGGGGCAGTGCCTTACACATGCTTGATTCCCACGCTACCCCCT
GCCTTGGGAGGTGTGTGGAATAAATTATTTTTGTTAAGGCA
```

FIGURE 212

MRRCNSGSGPPPSLLLLLLWLLAVPGANAAPRSALYSPSDPLTLLQADTVRGAVLGSRSAWAV

EFFASWCGHCIAFAPTWKALAEDVKAWRPALYLAALDCAEETNSAVCRDFNIPGFPTVRFFKA

FTKNGSGAVFPVAGADVQTLRERLIDALESHHDTWPPACPPLEPAKLEEIDGFFARNNEEYLA

LIFEKGGSYLGREVALDLSQHKGVAVRRVLNTEANVVRKFGVTDFPSCYLLFRNGSVSRVPVL

MESRSFYTAYLQRLSGLTREAAQTTVAPTTANKIAPTVWKLADRSKIYMADLESALHYILRIE

VGRFPVLEGQRLVALKKFVAVLAKYFPGRPLVQNFLHSVNEWLKRQKRNKIPYSFFKTALDDR

KEGAVLAKKVNWIGCQGSEPHFRGFPCSLWVLFHFLTVQAARQNVDHSQEAAKAKEVLPAIRG

YVHYFFGCRDCASHFEQMAAASMHRVGSPNAAVLWLWSSHNRVNARLAGAPSEDPQFPKVQWP

PRELCSACHNERLDVPVWDVEATLNFLKAHFSPSNIILDFPAAGSAARRDVQNVAAAPELAMG

ALELESRNSTLDPGKPEMMKSPTNTTPHVPAEGPEASRPPKLHPGLRAAPGQEPPEHMAELQR

NEQEQPLGQWHLSKRDTGAALLAESRAEKNRLWGPLEVRRVGRSSKQLVDIPEGQLEARAGRG

RGQWLQVLGGGFSYLDISLCVGLYSLSFMGLLAMYTYFQAKIRALKGHAGHPAA

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 705-728

N-glycosylation sites.

amino acids 130-134, 243-247, 575-579

Glycosaminoglycan attachment site.

amino acids 6-10

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 644-648

N-myristoylation sites.

amino acids 52-58, 56-62, 196-202, 381-387, 392-398, 448-454, 468-474, 684-690, 702-708

Cytochrome c family heme-binding site signature.

amino acids 509-515

Thioredoxin family proteins

amino acids 62-78

FIGURE 213

GCACGAGGCCGACTTCCAGACCATCTACAACTGCACGGCCTGGAACAGCTTCGGCTCCGACAC

TGAGATCATCCGGCTCAAGGAGCAAGGTTCGGAAATGAAGTCGGGAGCCGGGCTGGAAGCAGA

GTCTGTGCCG**ATG**GCCGTCATCATTGGGGTGGCCGTAGGAGCTGGTGTGGCCTTCCTCGTCCT

TATGGCAACCATCGTGGCGTTCTGCTGTGCCCGTTCCCAGAGAAATCTCAAAGGTGTTGTGTC

AGCCAAAAATGATATCCGAGTGGAAATTGTCCACAAGGAACCAGCCTCTGGTCGGGAGGGTGA

GGAGCACTCCACCATCAAGCAGCTGATGATGGACCGGGGTGAATTCCAGCAAGACTCAGTCCT

GAAACAGCTGGAGGTCCTCAAAGAAGAGGAGAAAGAGTTTCAGAACCTGAAGGACCCCACCAA

TGGCTACTACAGCGTCAACACCTTCAAAGAGCACCACTCAACCCCGACCATCTCCCTCTCCAG

CTGCCAGCCCGACCTGCGTCCTGCGGGTAAGCAGCGTGTGCCCACAGGCATGTCCTTCACCAA

CATCTACAGCACCCTGAGCGGCCAGGGCCGCCTCTACGACTACGGGCAGCGGTTTGTGCTGGG

CATGGGCAGCTCGTCCATCGAGCTTTGTGAGCGGGAGTTCCAGAGAGGCTCCCTCAGCGACAG

CAGCTCCTTCCTGGACACGCAGTGTGACAGCAGCGTCAGCAGCAGCGGCAAGCAGGATGGCTA

TGTGCAGTTCGACAAGGCCAGCAAGGCTTCTGCTTCCTCCTCCCACCACTCCCAGTCCTCGTC

CCAGAACTCTGACCCCAGTCGACCCCTGCAGCGGCGGATGCAGACTCACGTC**TAA**GGATCACA

CACCGCGGGTGGGGACGGGCCAGGGAAGAGGTCAGGGCACGTTCTGGTTGTCCAGGGACGAGG

GGTACTTTGCAGAGGACACCAGAATTGGCCACTTCCAGGACAGCCTCCCAGCGCCTCTGCCAC

TGCCTTCCTTCGAAGCTCTGATCAAGCACAAATCTGGGTCCCCAGGTGCTGTGTGCCAGAGGT

GGGCGGGTGGGGAGACAGACAGAGGCTGCGGCTGAGTGCGCTGTGCTTAGTGCTGGACACCCG

TGTCCCCGGCCCTTTCCTGGAGGCCCCTCTACCACCTGCTCTGCCCACAGGCACAAGTGGCAG

CTATAACTCTGCTTTCATGAAACTGCGGTCCACTCTCTGGTCTCTCTGTGGGCTCTACCCCTC

ACTGACCACAAGCTCTACCTACCCCTGTGCCTGTGCTCCCATACAGCCCTGGGGAGAAGGGGA

TGACGTCTTCCCAGCACTGAGCTGCCCCAGAAACCCCGGCTCCCCACTGCTGCTCATAGCCCA

TACCCTGGAGGCTGACAAGCCAGAAATGGCCTTGGCTAAAGGAGCCTCTCTCTCACCAGGCTG

GCCGGGAGCCCACCCCCAATTTGTTTGGTGTTTTGTGTCCATACTCTTGCAGTTCTGTCCTTG

GACTTGATGCCGCTGAACTCTGCGGTGGGACCGGTCCCGTCAGAGCCTGGTGTACTGGGGGGA

GGGAGGGAGGAGGGAGCCTGTGCTGACGGAGCACCTCGCCGGGTGTGCCCCTCCTGGGCTGTG

TGACCCCAGCCTCCCCACCCACCTCCTGCTTTGTGTACTCCTCCCCTCCCCCTCAGCACAATC

GGAGTTCATATAAGAAGTGCGGGAGCTTCTCTGGTCAGGGTTCTCTGAACACTTATGGAGAGA

GTGCTTCCTGGGAAGTGTGGCGTTTGAAGGGGCTGGAGGGCAGGTCTTTAAGATGGCGAGACT

GCCCTTCTCAGCTGATAAACACAAGAACGGCGATCCTGTCTTCAGTAAGGCTCCACGAGAAGA

GAGGAAGTATATCTACACCTCAACCCTCCTAGTCACCACCTGAAATAAATGTTAGGGAAAAAAAA

FIGURE 214

MAVIIGVAVGAGVAFLVLMATIVAFCCARSQRNLKGVVSAKNDIRVEIVHKEPASGREGEEHS

TIKQLMMDRGEFQQDSVLKQLEVLKEEEKEFQNLKDPTNGYYSVNTFKEHHSTPTISLSSCQP

DLRPAGKQRVPTGMSFTNIYSTLSGQGRLYDYGQRFVLGMGSSSIELCEREFQRGSLSDSSSF

LDTQCDSSVSSSGKQDGYVQFDKASKASASSSHHSQSSSQNSDPSRPLQRRMQTHV

Signal peptide:

amino acids 1-28

Glycosaminoglycan attachment site.

amino acids 150-154

N-myristoylation sites.

amino acids 6-12, 10-16, 36-42, 139-145, 165-171

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 114-125

FIGURE 215

```
CAGCCTTCCTCCCCCAGCCTGAGTGACTACTCTATTCCTTGGTCCCTGCTATTGTCGGGGACG
ATTGCATGGGCTACGCCAGGAAAGTAGGCTGGGTGACCGCAGGCCTGGTGATTGGGGCTGGCG
CCTGCTATTGCATTTATAGACTGACTAGGGGAAGAAAACAGAACAAGGAAAAAATGGCTGAGG
GTGGATCTGGGGATGTGGATGATGCTGGGGACTGTTCTGGGGCCAGGTATAATGACTGGTCTG
ATGATGATGACAGCAATGAGAGCAAGAGTATAGTATGGTACCCACCTTGGGCTCGGATTG
GGACTGAAGCTGGAACCAGAGCTAGGGCCAGGGCAAGGGCCAGGGCTACCCGGGCACGTCGGG
CTGTCCAGAAACGGGCTTCCCCCAATTCAGATGATACCGTTTTGTCCCCTCAAGAGCTACAAA
AGGTTCTTTGCTTGGTTGAGATGTCTGAAAGCCTTATATTCTTGAAGCAGCTTTAATTGCTC
TGGGTAACAATGCTGCTTATGCATTTAACAGAGATATTATTCGTGATCTGGGTGGTCTCCCAA
TTGTCGCAAAGATTCTCAATACTCGGGATCCCATAGTTAAGGAAAAGGCTTTAATTGTCCTGA
ATAACTTGAGTGTGAATGCTGAAAATCAGCGCAGGCTTAAAGTATACATGAATCAAGTGTGTG
ATGACACAATCACTTCTCGCTTGAACTCATCTGTGCAGCTTGCTGGACTGAGATTGCTTACAA
ATATGACTGTTACTAATGAGTATCAGCACATGCTTGCTAATTCCATTTCTGACTTTTTTCGTT
TATTTTCAGCGGGAAATGAAGAAACCAAACTTCAGGTTCTGAAACTCCTTTTGAATTTGGCTG
AAAATCCAGCCATGACTAGGGAACTGCTCAGGGCCCAAGTACCATCTTCACTGGGCTCCCTCT
TTAATAAGAAGGAGAACAAAGAAGTTATTCTTAAACTTCTGGTCATATTTGAGAACATAAATG
ATAATTTCAAATGGGAAGAAAATGAACCTACTCAGAATCAATTCGGTGAAGGTTCACTTTTTT
TCTTTTTAAAAGAATTTCAAGTGTGTGCTGATAAGGTTCTGGGAATAGAAAGTCACCATGATT
TTTTGGTGAAAGTAAAAGTTGGAAAATTCATGGCCAAACTTGCTGAACATATGTTCCCAAAGA
GCCAGGAATAACACCTTGATTTTGTAATTTAGAAGCAACACACATTGTAAACTATTCATTTTC
TCCACCTTGTTTATATGGTAAAGGAATCCTTTCAGCTGCCAGTTTTGAATAATGAATATCATA
TTGTATCATCAATGCTGATATTTAACTGAGTTGGTCTTTAGGTTTAAGATGGATAAATGAATA
TCACTACTTGTTCTGAAAACATGTTTGTTGCTTTTTATCTCGCTGCCTAGATTGAAATATTTT
GCTATTTCTTCTGCATAAGTGACAGTGAACCAATTCATCATGAGTAAGCTCCCTTCTGTCATT
TTCATTGATTTAATTTGTGTATCATCAATAAAATTGTATGTTAATGCTGGAAAGA
```

FIGURE 216

MGYARKVGWVTAGLVIGAGACYCIYRLTRGRKQNKEKMAEGGSGDVDDAGDCSGARYNDWSDD
DDDSNESKSIVWYPPWARIGTEAGTRARARARARATRARRAVQKRASPNSDDTVLSPQELQKV
LCLVEMSEKPYILEAALIALGNNAAYAFNRDIIRDLGGLPIVAKILNTRDPIVKEKALIVLNN
LSVNAENQRRLKVYMNQVCDDTITSRLNSSVQLAGLRLLTNMTVTNEYQHMLANSISDFFRLF
SAGNEETKLQVLKLLLNLAENPAMTRELLRAQVPSSLGSLFNKKENKEVILKLLVIFENINDN
FKWEENEPTQNQFGEGSLFFFLKEFQVCADKVLGIESHHDFLVKVKVGKFMAKLAEHMFPKSQE

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 68-72, 189-193, 217-221, 230-234

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-111

N-myristoylation sites.

amino acids 13-19, 17-23, 19-25, 54-60, 83-89, 147-153, 255-261, 290-296

Amidation site.

amino acids 29-33

FIGURE 217

GAGACACAAAGGCAGGCGGGATGCGGGAGCAGGCAAAGGGAAAGCGAAAGCCGCGCGCCCGGC

CGGTGACTGGGTGAAGGCGCCGCGCAGCTTTCCCGACGCCGGCTGTACCCGGACCTCCTGGTC

GAGCCTGGCGCGCCGCAGCC**ATG**GCCATCGCTCAACTGGCCACGGAGTACGTGTTCTCGGATT

TCTTGCTGAAGGAGCCCACGGAGCCCAAGTTCAAGGGGCTGCGACTGGAGCTGGCTGTGGACA

AGATGGTCACGTGCATTGCGGTGGGGCTGCCCCTGCTGCTCATCTCGCTGGCCTTCGCGCAGG

AGATCTCGATTGGTACACAGATAAGCTGTTTCTCTCCAAGTTCTTTCTCCTGGCGTCAGGCTG

CCTTTGTGGATTCATATTGCTGGGCGGCTGTTCAGCAGAAGAACTCACTGCAGAGCGAGTCTG

GAAACCTCCCACTGTGGCTGCATAAGTTTTTCCCCTACATCCTGCTGCTCTTTGCGATCCTCC

TGTACCTGCCCCCGCTGTTCTGGCGTTTCGCAGCTGCTCCTCATATTTGCTCAGACTTGAAGT

TTATCATGGAAGAACTTGACAAAGTTTACAACCGTGCAATTAAGGCTGCAAAGAGTGCGCGTG

ACCTTGACATGAGAGATGGAGCCTGCTCAGTTCCAGGTGTTACCGAGAACTTAGGGCAAAGTT

TGTGGGAGGTATCTGAAAGCCACTTCAAGTACCCAATTGTGGAGCAGTACTTGAAGACAAAGA

AAAATTCTAATAATTTAATCATCAAGTACATTAGCTGCCGCCTGCTGACACTCATCATTATAC

TGTTAGCGTGTATCTACCTGGGCTATTACTTCAGCCTCTCCTCACTCTCAGACGAGTTTGTGT

GCAGCATCAAATCAGGGATCCTGAGAAACGACAGCACCGTGCCCGATCAGTTTCAGTGCAAAC

TCATTGCCGTGGGCATCTTCCAGTTGCTCAGTGTCATTAACCTTGTGGTTTATGTCCTGCTGG

CTCCCGTGGTTGTCTACACGCTGTTTGTTCCATTCCGACAGAAGACAGATGTTCTCAAAGTGT

ACGAAATCCTCCCCACTTTTGATGTTCTGCATTTCAAATCTGAAGGGTACAACGATTTGAGCC

TCTACAATCTCTTCTTGGAGGAAAATATAAGTGAGGTCAAGTCATACAAGTGTCTTAAGGTAC

TGGAGAATATTAAGAGCAGTGGTCAGGGGATCGACCCAATGCTACTCCTGACAAACCTTGGCA

TGATCAAGATGGATGTTGTTGATGGCAAAACTCCCATGTCTGCAGAGATGAGAGAGGAGCAGG

GGAACCAGACGGCAGAGCTCCAAGGTATGAACATAGACAGTGAAACTAAAGCAAATAATGGAG

AGAAGAATGCCCGACAGAGACTTCTGGATTCTTCTTGC**TGA**TGATTTTTTTTCCTTGAGCTGT

AAATCTGTGACTTCTGCGACATGGGATTTAATTTGGCTAAAGCACCCCTGTTGGTTTCACAGC

TGGTTTGCAATAAATGGTTCTTGGTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 218

MAIAQLATEYVFSDFLLKEPTEPKFKGLRLELAVDKMVTCIAVGLPLLLISLAFAQEISIGTQ

ISCFSPSSFSWRQAAFVDSYCWAAVQQKNSLQSESGNLPLWLHKFFPYILLLFAILLYLPPLF

WRFAAAPHICSDLKFIMEELDKVYNRAIKAAKSARDLDMRDGACSVPGVTENLGQSLWEVSES

HFKYPIVEQYLKTKKNSNNLIIKYISCRLLTLIIILLACIYLGYYFSLSSLSDEFVCSIKSGI

LRNDSTVPDQFQCKLIAVGIFQLLSVINLVVYVLLAPVVVYTLFVPFRQKTDVLKVYEILPTF

DVLHFKSEGYNDLSLYNLFLEENISEVKSYKCLKVLENIKSSGQGIDPMLLLTNLGMIKMDVV

DGKTPMSAEMREEQGNQTAELQGMNIDSETKANNGEKNARQRLLDSSC

Transmembrane domains:

amino acids 37-55, 108-126, 216-232, 273-290

N-glycosylation sites.

amino acids 255-259, 338-342, 394-398

Glycosaminoglycan attachment site.

amino acids 357-361

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 203-207

N-myristoylation sites.

amino acids 61-67, 174-180, 251-257, 393-399

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 218-229

FIGURE 219

CTGTGAGTGACACACGCTGAGTGGGGTGAAGGGAA**ATG**CTGGTGAATTTCATTTTGAGGTGTG

GGTTGCTGTTAGTCACTCTGTCTCTTGCCATTGCCAAGCACAAGCAATCTTCCTTCACCAAAA

GTTGTTACCCAAGGGGAACATTGTCCCAAGCTGTTGACGCTCTCTATATCAAAGCAGCATGGC

TCAAAGCAACGATTCCAGAAGACCGCATAAAAAATATACGATTATTAAAAAAGAAAACAAAAA

AGCAGTTTATGAAAAACTGTCAATTTCAAGAACAGCTTCTGTCCTTCTTCATGGAAGACGTTT

TTGGTCAACTGCAATTGCAAGGCTGCAAGAAAATACGCTTTGTGGAGGACTTTCATAGCCTTA

GGCAGAAATTGAGCCACTGTATTTCCTGTGCTTCATCAGCTAGAGAGATGAAATCCATTACCA

GGATGAAAAGAATATTTTATAGGATTGGAAACAAAGGAATCTACAAAGCCATCAGTGAACTGG

ATATTCTTCTTTCCTGGATTAAAAAATTATTGGAAAGCAGTCAG**TAA**ACCAAAGCCAAGTACA

TTGATTTTACAGTTATTTTGAAATACAATAAGAACTGCTAGAAATATGTTTATAACAGTCTAT

TTCTTTTAAAAACTTTTTAACATAATACTGACGGCATGTTAGGTGATTCAGAATAGACAAGAA

GGATTTAGTAAATTAACGTTTTGGATATAAGTTGTCACTAATTTGCACATTTTCTGTGTTTTC

AAATAATGTTTCCATTCTGAACATGTTTTGTCATTCACAAGTACATTGTGTCAACTTAATTTA

AAGTATGTAACCTGAATTAACTCGTGTAATATTTGTGTGTGGAGTGGGATGTGGGGGGTGGAG

GGGGAATGACAGATTTCTGGAATGCAATGTAATGTTACTGAGACTTAAATAGATGTTATGTAT

ATGATTGTCTGTTTAAGTGTTTGAAAATTGTTAATTATGCCCAGTGTGAACTTAGTACTTAAC

ACATTTTGATTTTAATTAAATAAATTGGGTTTCCTTCTCAAAAAAAAAAAAAAAAAAAAAAAA

AAAAA

FIGURE 220

MLVNFILRCGLLLVTLSLAIAKHKQSSFTKSCYPRGTLSQAVDALYIKAAWLKATIPEDRIKN
IRLLKKKTKKQFMKNCQFQEQLLSFFMEDVFGQLQLQGCKKIRFVEDFHSLRQKLSHCISCAS
SAREMKSITRMKRIFYRIGNKGIYKAISELDILLSWIKKLLESSQ

Signal sequence:

amino acids 1-21

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 68-71

N-myristoylation site.

amino acids 148-153

Interleukin-10 proteins.

amino acids 58-94, 74-102, 128-170

FIGURE 221

GACCACGGCCCTGCGCCCCAGCCAGGCCTGAGGAC**ATG**AGGCGGCCGGCGGCGGTGCCGCTCC

TGCTGCTGCTGTGTTTTGGGTCTCAGAGGGCCAAGGCAGCAACAGCCTGTGGTCGCCCCAGGA

TGCTGAACCGAATGGTGGGCGGGCAGGACACGCAGGAGGGCGAGTGGCCCTGGCAAGTCAGCA

TCCAGCGCAACGGAAGCCACTTCTGCGGGGGCAGCCTCATCGCGGAGCAGTGGGTCCTGACGG

CTGCGCACTGCTTCCGCAACACCTCTGAGACGTCCCTGTACCAGGTCCTGCTGGGGGCAAGGC

AGCTAGTGCAGCCGGGACCACACGCTATGTATGCCCGGGTGAGGCAGGTGGAGAGCAACCCCC

TGTACCAGGGCACGGCCTCCAGCGCTGACGTGGCCCTGGTGGAGCTGGAGGCACCAGTGCCCT

TCACCAATTACATCCTCCCCGTGTGCCTGCCTGACCCCTCGGTGATCTTTGAGACGGGCATGA

ACTGCTGGGTCACTGGCTGGGGCAGCCCCAGTGAGGAAGACCTCCTGCCCGAACCGCGGATCC

TGCAGAAACTCGCTGTGCCCATCATCGACACACCCAAGTGCAACCTGCTCTACAGCAAAGACA

CCGAGTTTGGCTACCAACCCAAAACCATCAAGAATGACATGCTGTGCGCCGGCTTCGAGGAGG

GCAAGAAGGATGCCTGCAAGGGCGACTCGGGCGGCCCCCTGGTGTGCCTCGTGGGTCAGTCGT

GGCTGCAGGCGGGGGTGATCAGCTGGGGTGAGGGCTGTGCCCGCCAGAACCGCCCAGGTGTCT

ACATCCGTGTCACCGCCCACCACAACTGGATCCATCGGATCATCCCCAAACTGCAGTTCCAGC

CAGCGAGGTTGGGCGGCCAGAAG**TGA**GACCCCCGGGGCCAGGAGCCCCTTGAGCAGAGCTCTG

CACCCAGCCTGCCCGCCCACACCATCCTGCTGGTCCTCCCAGCGCTGCTGTTGCACCTGTGAG

CCCCACCAGACTCATTTGTAAATAGCGCTCCTTCCTCCCCTCTCAAATACCCTTATTTTATTT

ATGTTTCTCCCAATAAAAACCCAGCCTGTGTGCCAGCTGAAAAAAAAAAAAAAAAAAAA

FIGURE 222

MRRPAAVPLLLLLCFGSQRAKAATACGRPRMLNRMVGGQDTQEGEWPWQVSIQRNGSHFCGGS
LIAEQWVLTAAHCFRNTSETSLYQVLLGARQLVQPGPHAMYARVRQVESNPLYQGTASSADVA
LVELEAPVPFTNYILPVCLPDPSVIFETGMNCWVTGWGSPSEEDLLPEPRILQKLAVPIIDTP
KCNLLYSKDTEFGYQPKTIKNDMLCAGFEEGKKDACKGDSGGPLVCLVGQSWLQAGVISWGEG
CARQNRPGVYIRVTAHHNWIHRIIPKLQFQPARLGGQK

Important features of the protein:
Signal peptide:
amino acids 1-22

N-glycosylation sites.
amino acids 55-58, 79-82

Casein kinase II phosphorylation sites.
amino acids 121-124, 165-168, 167-170, 248-251

Tyrosine kinase phosphorylation sites.
amino acids 78-86, 197-203

N-myristoylation sites.
amino acids 16-21, 37-42, 56-61, 62-67, 118-123

Amidation site.
amino acids 219-222

Serine proteases, trypsin family, histidine active site.
amino acids 71-76

FIGURE 223

CAAG**ATG**TGGACAGCTCTTGTGCTCATTTGGATTTTCTCCTTGTCCTTATCTGAAAGCCATGC

GGCATCCAACGATCCACGCAACTTTGTCCCTAACAAAATGTGGAAGGGATTAGTCAAGAGGAA

TGCATCTGTGGAAACAGTTGATAATAAAACGTCTGAGGATGTAACCATGGCAGCAGCTTCTCC

TGTCACATTGACCAAAGGGACTTCGGCAGCCCACCTCAACTCTATGGAAGTCACAACAGAGGA

CACAAGCAGGACAGATGTGAGTGAACCAGCAACTTCAGGAGTTGCAGCTGATGGTGTGACCTC

CATTGCTCCCACGGCTGTGGCCTCCAGTACGACTGCGGCCTCCATTACGACTGCGGCCTCCAG

TATGACTGTGGCCTCCAGTGCTCCCACGACTGCAGCCTCCAGTACAACTGTGGCCTCCATTGC

TCCCACGACTGCAGCCTCCAGTATGACTGCGGCCTCCAGCACTCCCATGACACTTGCACTCCC

CGCGCCCACGTCCACTTCCACAGGGCGGACCCCGTCCACTACCGCCACTGGGCATCCATCTCT

CAGCACAGCCCTCGCACAAGTGCCAAAGAGCAGCGCGTTGCCAAGAACAGCAACCCTGGCCAC

ATTGGCCACACGTGCTCAGACTGTAGCGACCACAGCAAACACAAGCAGCCCCATGAGCACTCG

TCCAAGTCCTTCCAAGCACATGCCCAGTGACACCGCGGCAAGCCCTGTACCCCCTATGCGTCC

CCAAGCACAAGGTCCCATTAGCCAGGTGTCAGTGGACCAGCCTGTGGTTAACACAACAAATAA

ATCCACACCCATGCCCTCAAACACAACCCCAGAGCCCGCCCCCACCCCCACAGTGGTGACCAC

CACCAAGGCACAAGCCAGGGAGCCAACTGCCAGCCCAGTGCCAGTACCTCACACCAGCCCAAT

CCCTGAGATGGAGGCCATGTCCCCCACGACACAGCCAAGCCCCATGCCATATACCCAGAGGGC

CGCTGGGCCAGGCACATCCCAGGCACCGGAGCAGGTAGAGACTGAAGCCACACCAGGTACTGA

TTCCACTGGGCCAACACCCAGGAGCTCAGGGGGCACTAAGATGCCAGCCACGGACTCGTGCCA

GCCCAGCACCCAAGGCCAGTACATGGTGGTCACCACTGAGCCCCTCACCCAGGCCGTGGTAGA

CAAAACTCTCCTTCTGGTGGTGCTGTTACTCGGGGTGACCCTTTTCATCACAGTCTTGGTTTT

GTTTGCCCTGCAGGCCTATGAGAGCTACAAGAAGAAGGACTACACCCAGGTGGACTACTTAAT

CAACGGGATGTATGCGGACTCAGAAATG**TGA**GGGGGGCGGGGGCCTGGCGGGAGGCCTGGCCC

CTTCCTCGTCCTTTCCTTTTGCCTTTGAGACCAAACCAAGTGCTTCCAAATTCTTTTGGTGCA

ATTGAGGAGATATGCCAGATGCTTAAACACATTTAATTGCTGTCAGATTAATTCCATGATCAC

TAAAGAGTTGCTGCTTTTTTCATATTTATTTTTGTAAATGATTCTGTGCCCAGGAGCAGCTGG

GGGTTCCACCTCAGGGTGGGGCGGGCAGGACCCCGTCTCCCCAGGTGTCGGAGCCTGACCTGA

ATTAAAGTACTGACTGCTCGCCA

FIGURE 224

MWTALVLIWIFSLSLSESHAASNDPRNFVPNKMWKGLVKRNASVETVDNKTSEDVTMAAASPV
TLTKGTSAAHLNSMEVTTEDTSRTDVSEPATSGVAADGVTSIAPTAVASSTTAASITTAASSM
TVASSAPTTAASSTTVASIAPTTAASSMTAASSTPMTLALPAPTSTSTGRTPSTTATGHPSLS
TALAQVPKSSALPRTATLATLATRAQTVATTANTSSPMSTRPSPSKHMPSDTAASPVPPMRPQ
AQGPISQVSVDQPVVNTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAREPTASPVPVPHTSPIP
EMEAMSPTTQPSPMPYTQRAAGPGTSQAPEQVETEATPGTDSTGPTPRSSGGTKMPATDSCQP
STQGQYMVVTTEPLTQAVVDKTLLLVVLLLGVTLFITVLVLFALQAYESYKKKDYTQVDYLIN
GMYADSEM

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 396-420

N-glycosylation sites.

amino acids 41-44, 49-52, 222-225, 268-271, 271-274

Casein kinase II phosphorylation sites.

amino acids 14-17, 51-54, 80-83, 85-88, 280-283, 434-437

N-myristoylation sites.

amino acids 68-73, 354-359

Aldo/keto reductase family putative active site signature.

amino acids 195-210

FIGURE 225

GGAAGGCGCTCAAGGTGCGCGGCCCGGGGCGCGCTACTGGGGGCGCCCTCCGCGGTGGGCAGC

GCGCCAGGGATCGGCCTGGGCAGCCGCGGGGCGCGCGAAGGCTGCGCTTTCCCTACGGCCCCC

CTCGCTTCCTCCGGCACGGCGGCAACGGAGATTTCCTCTCGGGGAAACTACGCGGATCCTTTT

CGGGGATCCTCGCCCCGCCCCAGTTCTCCGCCCCCTCCCCTTTGCTGGGGCGCCTGGGCTGGC

CCGCGCAGGGGAGGAGGCTCTGGCAGCCTGGGCAGGGAGGCGGCGGGGGGCCGCGGAGCCGCT

GGCCATCGATTCTCCCCGCCATGTGACGCCGTCCTTAGCCCTGCGACCCCCAGCGCGTCCCGG

GCCTGCGCCTCCGCCCCGCCGCGCAGCGCACG**ATG**CTTCTGCCGGGACGCGCACGCCAACCGC

CGACGCCCCAGCCCGTGCAGCATCCCGGCCTCCGCCGGCAGGTAGAGCCGCCGGGGCAGCTCC

TGCGCCTCTTCTACTGCACTGTCCTGGTCTGCTCCAAAGAGATCTCAGCGCTCACCGACTTCT

CTGGTTACCTAACCAAACTCCTGCAAAACCACACCACCTATGCCTGTGATGGGGACTATTTGA

ATCTACAGTGCCCTCGGCATTCTACGATAAGTGTCCAATCGGCATTTTATGGGCAAGATTACC

AAATGTGTAGTTCCCAGAAGCCTGCCTCCCAGAGGGAAGACAGCTTAACCTGTGTGGCAGCCA

CCACCTTCCAGAAGGTGCTGGACGAATGCCAGAACCAGCGGGCCTGCCACCTCCTGGTCAATA

GCCGTGTTTTTGGACCTGACCTTTGTCCAGGAAGCAGTAAATACCTCCTGGTCTCCTTTAAAT

GCCAACCTAATGAATTAAAAAACAAAACCGTGTGTGAAGACCAGGAGCTGAAACTGCACTGCC

ATGAATCCAAGTTCCTCAACATCTACTCTGCGACCTACGGCAGGAGGACCCAGGAAAGGGACA

TCTGCTCCTCCAAGGCAGAGCGGCTCCCCCCTTTCGATTGCTTGTCTTACTCAGCTTTGCAAG

TCCTATCCCGAAGGTGCTATGGGAAGCAGAGATGCAAAATCATCGTCAACAATCACCATTTTG

GAAGCCCCTGTTTGCCAGGCGTGAAAAAATACCTCACTGTGACCTACGCATGTGTTCCCAAGA

ACATACTCACAGCGATTGATCCAGCCATTGCTAATCTAAAACCTTCTTTGAAGCAGAAAGATG

GTGAATATGGTATAAACTTCGACCCAAGCGGATCGAAGGTTCTGAGGAAAGATGGAATTCTTG

TTAGCAACTCTCTGGCAGCCTTTGCTTACATTAGAGCCCACCCAGAGAGAGCTGCCCTGCTGT

TCGTGTCCAGTGTCTGCATCGGCCTGGCCCTCACACTGTGCGCCCTGGTCATCAGAGAGTCCT

GTGCCAAGGACTTCCGCGACTTGCAGCTGGGGAGGGAGCAGCTGGTGCCAGGAAGTGACAAGG

TCGAGGAGGACAGCGAGGATGAAGAAGAGGAGGAGGACCCCTCTGAGTCTGATTTCCCAGGGG

AACTGTCGGGGTTCTGTAGGACTTCATATCCTATATACAGTTCCATAGAAGCTGCAGAGCTCG

CAGAAAGGATTGAGCGCAGGGAGCAAATCATTCAGGAAATATGGATGAACAGTGGTTTGGACA

CCTCGCTCCCAAGAAACATGGGCCAGTTCTAC**TGA**AAACCACATGCATCTTGATGCGATCGCA

CTTTCTGAAGAAGGAAGGATCCCAAATGCCCCTCCAGTTCTGGTTCACCTGTACCTTCTATGA

AGGAGAATTCGTCATGTCATTCAACACTCGTGAGGCCAGGAAGCTATTAAAGGGATGTTTCAA

GCTGTTTCTAGCACATTCCAAAATAAATGAGGAGGGAGGAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAA

FIGURE 226

MLLPGRARQPPTPQPVQHPGLRRQVEPPGQLLRLFYCTVLVCSKEISALTDFSGYLTKLLQNH
TTYACDGDYLNLQCPRHSTISVQSAFYGQDYQMCSSQKPASQREDSLTCVAATTFQKVLDECQ
NQRACHLLVNSRVFGPDLCPGSSKYLLVSFKCQPNELKNKTVCEDQELKLHCHESKFLNIYSA
TYGRRTQERDICSSKAERLPPFDCLSYSALQVLSRRCYGKQRCKIIVNNHHFGSPCLPGVKKY
LTVTYACVPKNILTAIDPAIANLKPSLKQKDGEYGINFDPSGSKVLRKDGILVSNSLAAFAYI
RAHPERAALLFVSSVCIGLALTLCALVIRESCAKDFRDLQLGREQLVPGSDKVEEDSEDEEEE
EDPSESDFPGELSGFCRTSYPIYSSIEAAELAERIERREQIIQEIWMNSGLDTSLPRNMGQFY

Transmembrane domains:

amino acids 32-49, 322-343

N-glycosylation sites.

amino acids 62-66, 165-169

Tyrosine kinase phosphorylation site.

amino acids 280-287

N-myristoylation site.

amino acids 302-308, 333-339, 428-434

Amidation site.

amino acids 191-195

FIGURE 227

GGCACGAGGTGGAAGGGCTTTTACAAACAGATTGCTGGCCCCACCCCCCAGAATTTCTCATCA

GGAGTGGGCAAGACCAATCATTTGCATTTCTGACAAGTTCCCAGGAGCTGCAGCTGCTGGCCC

TGGAACCACACTTTGAGAACCACTGCTTTAGACCAAACACCAAAGGAAGATGCAGCCACCCTC

CTTTACATGTCACAACGCTCAGGGTCCATGAGTACCTCAGGCTGTCCAGCTGAGCTCCACCTG

CAGCAGCCGAGATTCCCGACTCGCTCCACCATTGGGGGCTAGGAGTGAAGCGTGTCACC**ATG**G

TCAGCTCATGGCCAGCCAGGAAAGCCTCTCTGCTGTGCGTCTGTGCAGTTCTTGTTCTTCCCT

GGAGGACTCTTGGATCGCCTGTGATCTTGGCCAGGAGACCAGGTGCCTGGGTCCCTTCCTGGA

AGGGGACAAGTTACACACCCCAGCCCCATTTTCCCACCAACTTCTACATGCCTTGGGAGAACC

TTCTACATGTTGGCTGCCCCCTTCCCCTATTTCAGCAGTGCCCAGTCCTGCTTATAAACCTGA

GGCCTGCTCCCCATACCTTCCCTGTGCAAGTGCCAGCCGTTATTCCAGGCAGCCCAATGTTGT

TGAGGCCAGATGGATTCCTGGAAGCAGCTGGCCCATGGATG**TGA**GTCATCACAGTATTCTAGA

AACAGAGAAGAGGTCTTAACCTAATGCGCATAGAGAAATTGTTCTCATTGTAAACATACCCCT

GTCCTTAGCTGATCTAGGTGGAAGCCCAGCTTCATGTGCTAGGGGGCATGATAATGATAATAA

AGGAATTGTATCTAGGACTAA

FIGURE 228

MVSSWPARKASLLCVCAVLVLPWRTLGSPVILARRPGAWVPSWKGTSYTPQPHFPTNFYMPWE
NLLHVGCPLPLFQQCPVLLINLRPAPHTFPVQVPAVIPGSPMLLRPDGFLEAAGPWM

Signal peptide:
amino acids 1-27

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 8-12

FIGURE 229

```
GGGAAGGGATGCAAGGAAGCCCTCCGGCGCTGCGCTCCGAGGCGGGAGACAGCGTCCCGCTGA
AAATGTGTGTCTGACATGCAAGCTCAGTGGGGCAGAGACCCGTGGATTGCTGTGCCCTGCCCT
CCGGACCTGGATCATGAAGGTGTTGGGAAGAAGCTTCTTCTGGGTGCTGTTTCCCGTCCTTCC
CTGGGCGGTGCAGGCTGTGGAGCACGAGGAGGTGGCGCAGCGTGTGATCAAACTGCACCGCGG
GCGAGGGGTGGCTGCCATGCAGAGCCGGCAGTGGGTCCGGGACAGCTGCAGGAAGCTCTCAGG
GCTTCTCCGCCAGAAGAATGCAGTTCTGAACAAACTGAAAACTGCAATTGGAGCAGTGGAGAA
AGACGTGGGCCTGTCGGATGAAGAGAAACTGTTTCAGGTGCACACGTTTGAAATTTTCCAGAA
AGAGCTGAATGAAAGTGAAAATTCCGTTTTCCAAGCTGTCTACGGACTGCAGAGAGCCCTGCA
GGGGGATTACAAAGATGTCGTGAACATGAAGGAGAGCAGCCGGCAGCGCCTGGAGGCCCTGAG
AGAGGCTGCAATAAAGGAAGAAACAGAATATATGGAACTTCTGGCAGCAGAAAAACATCAAGT
TGAAGCCCTTAAAAATATGCAACATCAAAACCAAAGTTTATCCATGCTTGACGAGATTCTTGA
AGATGTAAGAAAGGCAGCGGATCGTCTGGAGGAAGAGATAGAGGAACATGCTTTTGACGACAA
TAAATCAGTCAAGGGGGTCAATTTTGAGGCAGTTCTGAGGGTGGAGGAAGAAGAGGCCAATTC
TAAGCAAAATATAACAAACGAGAAGTGGAGGATGACTTGGGTCTTAGCATGCTGATTGACTC
CCAGAACAACCAGTATATTTTGACCAAGCCCAGAGATTCAACCATCCCACGTGCAGATCACCA
CTTTATAAAGGACATTGTTACCATAGGAATGCTGTCCTTGCCTTGTGGCTGGCTATGTACAGC
CATAGGATTGCCTACAATGTTTGGTTATATTATTTGTGGTGTACTTCTGGGACCTTCAGGACT
AAATAGTATTAAGTCTATTGTGCAAGTGGAGACATTAGGAGAATTTGGGGTGTTTTTTACTCT
TTTTCTTGTTGGCTTAGAATTTTCTCCAGAAAAGCTAAGAAAGGTGTGGAAGATTTCCTTACA
AGGGCCGTGTTACATGACACTGTTAATGATTGCATTTGGCTTGCTGTGGGGGCATCTCTTGCG
GATCAAACCCACGCAGAGCGTCTTCATTTCCACGTGTCTGTCCTTGTCAAGCACACCCCTCGT
GTCCAGGTTCCTCATGGGCAGTGCTCGGGGTGACAAAGAAGGCGACATTGACTACAGCACCGT
GCTCCTCGGCATGCTGGTGACGCAGGACGTGCAGCTCGGGCTCTTCATGGCCGTCATGCCGAC
TCTCATACAGGCGGGCGCCAGTGCATCTTCTAGCATTGTCGTGGAAGTTCTCCGAATCCTGGT
TTTGATTGGTCAGATTCTTTTTTCACTAGCGGCGGTTTTTCTTTTATGTCTTGTTATAAAGAA
GTATCTCATTGGACCCTATTATCGGAAGCTGCACATGGAAAGCAAGGGGAACAAAGAAATCCT
GATCTTGGGAATATCTGCCTTTATCTTCTTAATGTTAACGGTCACGGAGCTGCTGGACGTCTC
CATGGAGCTGGGCTGTTTCCTGGCTGGAGCGCTCGTCTCCTCTCAGGGCCCCGTGGTCACCGA
GGAGATCGCCACCTCCATCGAACCCATCCGCGACTTCCTGGCCATCGTTTTCTTCGCCTCCAT
AGGGCTCCACGTGTTCCCCACGTTTGTGGCGTACGAGCTCACGGTGCTGGTGTTCCTCACCTT
GTCAGTGGTGGTGATGAAGTTTCTCCTGGCGGCGCTGGTCCTGTCTCTCATTCTGCCGAGGAG
CAGCCAGTACATCAAGTGGATCGTCTCTGCGGGGCTTGCCCAGGTCAGCGAGTTTTCCTTTGT
CCTGGGGAGCCGGGCGCGAAGAGCGGGCGTCATCTCTCGGGAGGTGTACCTCCTTATACTGAG
TGTGACCACGCTCAGCCTCTTGCTCGCCCCGGTGCTGTGGAGCTGCAATCACGAGGTGTGT
GCCCAGACCGGAGAGACGGTCCAGCCTCTGATGGCTCGGAGATGATGGACCGTGGAAGGGAAG
CGTCTGTGGGGAGTGAGCGCTTAGATGGCCAGCAGCTGCTCCTTCTGGGAAGCTCGCACCTTG
GCAACAGAACAGCCCTCTAGCAGAGCGTCAGTGCAGTCGTGTTATCCCGGCTTTTACAGAATA
TTCTTGTCCTATTTTAGAATTTTCCGGAGTAGTTTATTTGCAGTCTGTTGATTATGTGCAGTA
GACCCGGGACACTGCGTTTTACCGATCACCTTGAATGTGGTGCCTGGATGTGCCTTTTTTTTT
TTTCCCTGAAATTATTATTAATTTTCTATTGTGAGTTCATCAGTTCATAGTTTTTTTAGTAAA
GAAGCAAAATTAAAAGGCTTTTAAAAATGTACAACTTCAGAATTATAATCTGTTAGTCAAATA
TTTGTTATTAAACATTTCTGTAATATGAAGTTGTAATCCTGGCCGTGAGCTTGGAAGCTTACT
TTTGATTCTTAAAGCCTATGTTTTCTAAAATGAGACAAATACGGATGTCTATTTGCCTTTTAT
TGTAACTTTTAAATGAAATAATTTCATGTCAATTTCTATTAGATATATCACTTAAAATATTTG
GTTTTAAATCACAAGAATATGTATTCTTTAATAAAGATAATTTATGATCATGGTAAAAAAAAAA
```

FIGURE 230

MKVLGRSFFWVLFPVLPWAVQAVEHEEVAQRVIKLHRGRGVAAMQSRQWVRDSCRKLSGLLRQ
KNAVLNKLKTAIGAVEKDVGLSDEEKLFQVHTFEIFQKELNESENSVFQAVYGLQRALQGDYK
DVVNMKESSRQRLEALREAAIKEETEYMELLAAEKHQVEALKNMQHQNQSLSMLDEILEDVRK
AADRLEEEIEEHAFDDNKSVKGVNFEAVLRVEEEEANSKQNITKREVEDDLGLSMLIDSQNNQ
YILTKPRDSTIPRADHHFIKDIVTIGMLSLPCGWLCTAIGLPTMFGYIICGVLLGPSGLNSIK
SIVQVETLGEFGVFFTLFLVGLEFSPEKLRKVWKISLQGPCYMTLLMIAFGLLWGHLLRIKPT
QSVFISTCLSLSSTPLVSRFLMGSARGDKEGDIDYSTVLLGMLVTQDVQLGLFMAVMPTLIQA
GASASSSIVVEVLRILVLIGQILFSLAAVFLLCLVIKKYLIGPYYRKLHMESKGNKEILILGI
SAFIFLMLTVTELLDVSMELGCFLAGALVSSQGPVVTEEIATSIEPIRDFLAIVFFASIGLHV
FPTFVAYELTVLVFLTLSVVVMKFLLAALVLSLILPRSSQYIKWIVSAGLAQVSEFSFVLGSR
ARRAGVISREVYLLILSVTTLSLLLAPVLWRAAITRCVPRPERRSSL

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 282-304, 322-337, 354-370, 379-395, 445-474, 501-520, 576-598, 641-660

N-glycosylation sites.

amino acids 104-108, 174-178, 206-210, 230-234

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 55-59, 673-677

Tyrosine kinase phosphorylation site.

amino acids 407-414

N-myristoylation sites.

amino acids 116-122, 327-333, 366-372, 401-407, 419-425, 429-435, 442-448, 525-531, 530-536

Cell attachment sequence.

amino acids 404-407

FIGURE 231

GAGAAAAACAACAGGAAGCAGCTTACAAACTCGGTGAACAACTGAGGGAACCAAACCAGAGAC

GCGCTGAACAGAGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTG

GTGCAAGGCGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAG**ATG**CTGGGGAGCAGAGCTG

TAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCC

CTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACATC

CACTAGTGGGACACATGGATCTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCC

ATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGC

AAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGG

AGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCC

AACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCC

AGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTG

TAGCCGCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCC**TAA**AGGCAGCAGCTCAAGG

ATGGCACTCAGATCTCCATGGCCCAGCAAGGCCAAGATAAATCTACCACCCCAGGCACCTGTG

AGCCAACAGGTTAATTAGTCCATTAATTTTAGTGGGACCTGCATATGTTGAAAATTACCAATA

CTGACTGACATGTGATGCTGACCTATGATAAGGTTGAGTATTTATTAGATGGGAAGGGAAATT

TGGGGATTATTTATCCTCCTGGGGACAGTTTGGGGAGGATTATTTATTGTATTTATATTGAAT

TATGTACTTTTTTCAATAAAGTCTTATTTTTGTGGCTAAAAAAAAAAAA

FIGURE 232

MLGSRAVMLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLAWSAHPLVGHMDLREEGDEE
TTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEKLLGSDIFTGEPSLLPDSPVGQLHA
SLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLLLRFKILRSLQAFVAVAARVFAHGAATLSP

Important features of the protein:

Signal peptide:

amino acids 1-21

Casein kinase II phosphorylation site.

amino acids 64-67

N-myristoylation sites.

amino acids 25-30, 81-86, 122-127

FIGURE 233

```
CCCACGCGTCCGGCCCTGTAACCAAGATACTGACTGAACATGGCTGGCGGACTCAGGCTGGGGTCTGCAGTGCAG
CATTAATGGGCCGCTGACATGAATATGGAGTAGTTTTCTCTAGCAAAGAGTAATGTGGGCCATGGAGTCAGGCCA
CCTCCTCTGGGCTCTGCTGTTCATGCAGTCCTTGTGGCCTCAACTGACTGATGGAGCCACTCGAGTCTACTACCT
GGGCATCCGGGATGTGCAGTGGAACTATGCTCCCAAGGGAAGAAATGTCATCACGAACCAGCCTCTGGACAGTGA
CATAGTGGCTTCCAGCTTCTTAAAGTCTGACAAGAACCGGATAGGGGGAACCTACAAGAAGACCATCTATAAAGA
ATACAAGGATGACTCATACACAGATGAAGTGGCCCAGCCTGCCTGGTTGGGCTTCCTGGGGCCAGTGTTGCAGGC
TGAAGTGGGGGATGTCATTCTTATTCACCTGAAGAATTTTGCCACTCGTCCCTATACCATCCACCCTCATGGTGT
CTTCTACGAGAAGGACTCTGAAGGTTCCCTATACCCAGATGGCTCCTCTGGGCCACTGAAAGCTGATGACTCTGT
TCCCCGGGGGGCAGCCATATCTACAACTGGACCATTCCAGAAGGCCATGCACCCACCGATGCTGACCCAGCGTG
CCTCACCTGGATCTACCATTCTCATGTAGATGCTCCACGAGACATTGCAACTGGCCTAATTGGGCCTCTCATCAC
CTGTAAAAGAGGAGCCCTGGATGGGAACTCCCCTCCTCAACGCCAGGATGTAGACCATGATTTCTTCCTCCTCTT
CAGTGTGGTAGATGAGAACCTCAGCTGGCATCTCAATGAGAACATTGCCACTTACTGCTCAGATCCTGCTTCAGT
GGACAAAGAAGATGAGACATTTCAGGAGAGCAATAGGATGCATGCAATCAATGGCTTTGTTTTTGGGAATTTACC
TGAGCTGAACATGTGTGCACAGAAACGTGTGGCCTGGCACTTGTTTGGCATGGGCAATGAAATTGATGTCCACAC
AGCATTTTTCCATGGACAGATGCTGACTACCCGTGGACACCACACTGATGTGGCTAACATCTTTCCAGCCACCTT
TGTGACTGCTGAGATGGTGCCCTGGGAACCTGGTACCTGGTTAATTAGCTGCCAAGTGAACAGTCACTTTCGAGA
TGGCATGCAGGCACTCTACAAGGTCAAGTCTTGCTCCATGGCCCCTCCTGTGGACCTGCTCACAGGCAAAGTTCG
ACAGTACTTCATTGAGGCCCATGAGATTCAATGGGACTATGGCCCGATGGGGCATGATGGGAGTACTGGGAAGAA
TTTGAGAGAGCCAGGCAGTATCTCAGATAAGTTTTTCCAGAAGAGCTCCAGCCGAATTGGGGGCACTTACTGGAA
AGTGCGATATGAAGCCTTTCAAGATGAGACATTCCAAGAGAAGATGCATTTGGAGGAAGATAGGCATCTTGGAAT
CCTGGGGCCAGTGATCCGGGCTGAGGTGGGTGACACCATTCAGGTGGTCTTCTACAACCGTGCCTCCCAGCCATT
CAGCATGCAGCCCCATGGGGTCTTTTATGAGAAAGACTATGAAGGCACTGTGTACAATGATGGCTCATCTTACCC
TGGCTTGGTTGCCAAGCCCTTTGAGAAAGTAACATACCGCTGGACAGTCCCCCCTCATGCCGGTCCCACTGCTCA
GGATCCTGCTTGTCTCACTTGGATGTACTTCTCTGCTGCAGATCCCATAAGAGACACAAATTCTGGCCTGGTGGG
CCCGCTGCTGGTGTGCAGGGCTGGTGCCTTGGGTGCAGATGGCAAGCAGAAAGGGGTGGATAAAGAATTCTTTCT
TCTCTTCACTGTGTTGGATGAGAACAAGAGCTGGTACAGCAATGCCAATCAAGCAGCTGCTATGTTGGATTTCCG
ACTGCTTTCAGAGGATATTGAGGGCTTCCAAGACTCCAATCGGATGCATGCCATTAATGGGTTTCTGTTCTCTAA
CCTGCCCAGGCTGGACATGTGCAAGGGTGACACAGTGGCCTGGCACCTGCTCGGCCTGGGCACAGAGACTGATGT
GCATGGAGTCATGTTCCAGGGCAACACTGTGCAGCTTCAGGGCATGAGGAAGGGTGCAGCTATGCTCTTTCCTCA
TACCTTTGTCATGGCCATCATGCAGCCTGACAACCTTGGGACATTTGAGATTTATTGCCAGGCAGGCAGCCATCG
AGAAGCAGGGATGAGGGCAATCTATAATGTCTCCCAGTGTCCTGGCCACCAAGCCACCCCTCGCCAACGCTACCA
AGCTGCAAGAATCTACTATATCATGGCAGAAGAAGTAGAGTGGGACTATTGCCCTGACCGGAGCTGGGAACGGGA
ATGGCACAACCAGTCTGAGAAGGACAGTTATGGTTACATTTTCCTGAGCAACAAGGATGGGCTCCTGGGTTCCAG
ATACAAGAAAGCTGTATTCAGGGAATACACTGATGGTACATTCAGGATCCCTCGGCCAAGGACTGGACCAGAAGA
ACACTTGGGAATCTTGGGTCCACTTATCAAAGGTGAAGTTGGTGATATCCTGACTGTGGTATTCAAGAATAATGC
CAGCCGCCCCTACTCTGTGCATGCTCATGGAGTGCTAGAATCTACTACTGTCTGGCCACTGGCTGCTGAGCCTGG
TGAGGTGGTCACTTATCAGTGGAACATCCCAGAGAGGTCTGGCCCTGGGCCCAATGACTCTGCTTGTGTTTCCTG
GATCTATTATTCTGCAGTGGATCCCATCAAGGACATGTATAGTGGCCTGGTGGGGCCCTTGGCTATCTGCCAAAA
GGGCATCCTGGAGCCCCATGGAGGACGGAGTGACATGGATCGGGAATTTGCATTGTTGTTCTTGATTTTTGATGA
AAATAAGTCTTGGTATTTGGAGGAAAATGTGGCAACCCATGGGTCCCAGGATCCAGGCAGTATTAACCTACAGGA
TGAAACTTTCTTGGAGAGCAATAAAATGCATGCAATCAATGGGAAACTCTATGCCAACCTTAGGGGTCTTACCAT
GTACCAAGGAGAACGAGTGGCCTGGTACATGCTGGCCATGGGCCAAGATGTGGATCTACACACCATCCACTTTCA
TGCAGAGAGCTTCCTCTATCGGAATGGCGAGAACTACCGGGCAGATGTGGTGGATCTGTTCCCAGGGACTTTTGA
GGTTGTGGAGATGGTGGCCAGCAACCCTGGGACATGGCTGATGCACTGCCATGTGACTGACCATGTCCATGCTGG
CATGGAGACCCTCTTCACTGTTTTTTCTCGAACAGAACACTTAAGCCCTCTCACCGTCATCACCAAAGAGACTGA
AAAAGTGCCCCCCAGAGACATTGAAGAAGGCAATGTGAAGATGCTGGGCATGCAGATCCCCATAAAGAATGTTGA
GATGCTGGCCTCTGTTTTTGGTTGCCATTAGTGTCACCCTTCTGCTCGTTGTTCTGGCTCTTGGTGGAGTGGTTTG
GTACCAACATCGACAGAGAAAGCTACGACGCAATAGGAGGTCCATCCTGGATGACAGCTTCAAGCTTCTGTCTTT
CAAACAGTAACATCTGGAGCCTGGAGATATCCTCAGGAAGCACATCTGTAGTGCACTCCCAGCAGGCCATGGACT
AGTCACTAACCCCACACTCAAAGGGCATGGGTGGTGGAGAAGCAGAAGGAGCAATCAAGCTTATCTGGATATTT
CTTTCTTTATTTATTTTACATGGAAATAATATGATTTCACTTTTTCTTTAGTTTCTTTGCTCTACGTGGGCACCT
GGCACTAAGGGAGTACCTTATTATCCTACATCGCAAATTTCAACAGCTACATTATATTTCCTTCTGACACTTGGA
AGGTATTGAAATTTCTAGAAATGTATCCTTCTCACAAAGTAGAGACCAAGAGAAAAACTCATTGATTGGGTTTCT
ACTTCTTTCAAGGACTCAGGAAATTTCACTTTGAACTGAGGCCAAGTGAGCTGTTAAGATAACCCACACTTAAAC
TAAAGGCTAAGAATATAGGCTTGATGGGAAATTGAAGGTAGGCTGAGTATTGGGAATCCAAATTGAATTTTGATT
CTCCTTGGCAGTGAACTACTTTGAAGAAGTGGTCAATGGGTTGTTGCTGCCATGAGCATGTACAACCTCTGGAGC
TAGAAGCTCCTCAGGAAAGCCAGTTCTCCAAGTTCTTAACCTGTGGCACTGAAAGGAATGTTGAGTTACCTCTTC
ATGTTTTAGACAGCAAACCCTATCCATTAAAGTACTTGTTAGACCAAAAAAAAAAAAA
```

FIGURE 234

MWAMESGHLLWALLFMQSLWPQLTDGATRVYYLGIRDVQWNYAPKGRNVITNQPLDSDIVASS
FLKSDKNRIGGTYKKTIYKEYKDDSYTDEVAQPAWLGFLGPVLQAEVGDVILIHLKNFATRPY
TIHPHGVFYEKDSEGSLYPDGSSGPLKADDSVPPGGSHIYNWTIPEGHAPTDADPACLTWIYH
SHVDAPRDIATGLIGPLITCKRGALDGNSPPQRQDVDHDFFLLFSVVDENLSWHLNENIATYC
SDPASVDKEDETFQESNRMHAINGFVFGNLPELNMCAQKRVAWHLFGMGNEIDVHTAFFHGQM
LTTRGHHTDVANIFPATFVTAEMVPWEPGTWLISCQVNSHFRDGMQALYKVKSCSMAPPVDLL
TGKVRQYFIEAHEIQWDYGPMGHDGSTGKNLREPGSISDKFFQKSSSRIGGTYWKVRYEAFQD
ETFQEKMHLEEDRHLGILGPVIRAEVGDTIQVVFYNRASQPFSMQPHGVFYEKDYEGTVYNDG
SSYPGLVAKPFEKVTYRWTVPPHAGPTAQDPACLTWMYFSAADPIRDTNSGLVGPLLVCRAGA
LGADGKQKGVDKEFFLLFTVLDENKSWYSNANQAAAMLDFRLLSEDIEGFQDSNRMHAINGFL
FSNLPRLDMCKGDTVAWHLLGLGTETDVHGVMFQGNTVQLQGMRKGAAMLFPHTFVMAIMQPD
NLGTFEIYCQAGSHREAGMRAIYNVSQCPGHQATPRQRYQAARIYYIMAEEVEWDYCPDRSWE
REWHNQSEKDSYGYIFLSNKDGLLGSRYKKAVFREYTDGTFRIPRPRTGPEEHLGILGPLIKG
EVGDILTVVFKNNASRPYSVHAGVLESTTVWPLAAEPGEVVTYQWNIPERSGPGPNDSACVS
WIYYSAVDPIKDMYSGLVGPLAICQKGILEPHGGRSDMDREFALLFLIFDENKSWYLEENVAT
HGSQDPGSINLQDETFLESNKMHAINGKLYANLRGLTMYQGERVAWYMLAMGQDVDLHTIHFH
AESFLYRNGENYRADVVDLFPGTFEVVEMVASNPGTWLMHCHVTDHVHAGMETLFTVFSRTEH
LSPLTVITKETEKVPPRDIEEGNVKMLGMQIPIKNVEMLASVLVAISVTLLLVVLALGGVVWY
QHRQRKLRRNRRSILDDSFKLLSFKQ

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 1109-1130

N-glycosylation sites.

amino acids 167-171, 239-243, 591-595, 717-721, 761-765, 832-836, 876-880, 934-938

Glycosaminoglycan attachment site.

amino acids 871-875

Tyrosine kinase phosphorylation sites.

amino acids 82-90, 137-145, 494-502, 513-521

N-myristoylation sites.

amino acids 212-218, 313-319, 498-504, 566-572, 672-678, 778-784, 843-849

Multicopper oxidases signature 1.

amino acids 344-365, 696-717, 1043-1064

Multicopper oxidases signature 2.

amino acids 1048-1060

FIGURE 235

```
GGAAAGAGTGCTGGTACTACAACCAGGAAGTGACAGATAATGTGCTTTAAACTACATTAGAAAAGCTTCTCATAG
CAAAACTGAGAGATTGAAGCAGTGATTATTTTTACATAGTTGTCATTAAATATTTGGAGCTCTGCTGTGCATAGA
GATGGCAACATACTTAGAATACACAGCTTTCTGGGCCAGAAATTGATCTTCTGACTTTTGAGCCTTATCTGATTA
CTGCTTGGTTCATCTTTATTTTGTTAAACTACTCTGTAGGCTGAAAGGGAGAGACTCTCCTTGGTTTGCAGAGCC
TGACTAGACAGGAATTCTGGCAACTGCTCCAGCAGAACTATGGCACTGAGCTAGGTTTAAATGCTGAGGAG**ATG**G
AAAACTTGTCACTGTCGATTGAGGATGTGCAGCCAAGAAGTCCAGGAAGAAGCAGCTTGGATGACTCTGGGGAGA
GAGATGAAAAATTATCCAAGTCAATCAGTTTTACCAGTGAATCAATTAGTCGGGTTTCAGAAACAGAGTCATTCG
ATGGAAATTCATCAAAAGGAGGATTAGGCAAAGAGGAGTCCCAAAATGAGAAACAGACCAAAAAGAGTCTCTTAC
CAACTTTGGAAAAGAAGTTAACTAGAGTGCCATCAAAGTCACTGGACTTGAATAAAAATGAATATCTTTCTCTGG
ACAAAAGCAGCACTTCAGATTCTGTTGATGAAGAAAATGTTCCTGAGAAAGATCTTCATGGAAGACTTTTTATCA
ACCGTATTTTTCATATCAGTGCTGACAGAATGTTTGAATTGCTCTTTACCAGTTCACGCTTTATGCAGAAATTTG
CCAGTTCTAGAAATATAATAGATGTAGTATCTACCCCTTGGACTGCAGAACTTGGAGGTGATCAGCTGAGAACGA
TGACCTACACTATAGTCCTTAATAGTCCACTTACTGGAAAATGCACTGCTGCCACTGAAAAGCAGACACTGTATA
AAGAAAGTCGGGAAGCACGATTTTATTTGGTAGATTCAGAAGTACTGACACATGATGTCCCCTACCATGATTACT
TCTATACCGTGAACAGATACTGTATCATCCGATCTTCAAAACAGAAATGCAGGCTAAGAGTTTCCACAGATTTGA
AATACAGAAAACAGCCATGGGGCCTTGTCAAATCTTTAATTGAAAGAATTCCTGGAGTTCTTTGGAGGACTATT
TCAAACAGCTTGAATCAGATTTGTTAATTGAAGAATCTGTATTAAATCAGGCCATTGAAGACCCTGGAAAACTTA
CTGGCCTACGAAGGAGAAGGCGAACCTTCAACCGAACAGCAGAAACAGTTCCTAAACTTTCCTCTCAGCATTCCT
CTGGAGATGTGGGCTTAGGTGCCAAAGGGGATATTACAGGAAAGAAAAAGGAAATGGAAAACTATAACGTCACTC
TTATTGTGGTAATGAGTATTTTTGTGTTGTTATTAGTTTTGTTGAATGTGACACTGTTTCTGAAGCTGTCAAAGA
TAGAACATGCTGCTCAGTCCTTTTACCGTCTCCGCCTCCAAGAAGAGAAATCTTTAAATTTAGCCTCTGATATGG
TGTCAAGAGCAGAAACTATTCAGAAGAATAAAGATCAGGCCCATCGTTTAAAGGGAGTGCTCCGAGACTCCATAG
TGATGCTTGAACAGCTGAAGAGCTCACTCATTATGCTTCAGAAAACGTTTGATCTACTAAATAAGAATAAGACTG
GCATGGCTGTTGAAAGC**TAG**TGATCTGAAGGACTAAAACCGCAGAGATACTTGGAACTTAAAGAAAATACCTGGA
AGAAAACCAGACGAATGAAGGATTTTGGCATAGAACATTTCTATGTTTTTTCATTATTGAGATTTCTAATATGAA
CATTTCTTTCAGTAACATTTATTTGATAATTAGTTTCTGCTGGCCTTAATAATCCATCCTTTCACTTCTTATAGA
TATTTTTAAGCTGTGAATTTCTTCAGTGAACCATGAAATATATTATAGAACTGAATTTCTCTGATACAAAAAGAA
AATGACACACCCTGAATTGAGTGGTATGGTCTCATTTCTACAGTGAAGTCTGATGCTTTGTTAGCACAGAATCCG
TACATGTCCAATAGGTCGCTTTTGTAACTGAGATAAGACCAAGAGGATAAACAGGACAATATAAGAAGAAACCTC
TATGTCATTACTGATTTTAAAGGTTCTGTTTTCAGGCATATAACATTTCCAGGTTTGTGTACTGTAAAGATTATA
ATGTCTTCATTTATTTAGCATGCAAATTTAATAGTCAAACTTTTTGAATCTGCATGTTGATGATGATTATCAGAA
AGGGTCTTCTGCCATGCTGTATCTTTATGAAAGAAATAGTTGTTTTTTCTTAAGGTAACTATCAGAGGTGGGATT
ATCTTGCCTCCTCACTTAGAATACCAACAGTCAAAAGGAAGAACCATCCTCTGAGTTTTAAAAACCAGAAGGTTA
TGTTAAAATCTGGGCATTTAGTGACAGATCAAATGCATACTTGAACTAAGATTGGCTTCAGCTTAGCAGTCTTTC
ATGGTGGAAGTGACACATCTGGTTGAAAATAATTTGTGTATTTTCAGTAACCATGTATGGCTTCCTTCTTTATGT
ATGTGTGTGACTTGTTTTAATTGGTAAGTTATAAGCCAGACATAGATTTTAGCTCTTTAATAAAAACTTCAGGGG
CACGTATGTCCCAGTACAAGTGTACTGACTATCAAGTTTTAACTCAGATGCAAGCTTTGGCTCTTTCATAAAAAG
TTTTTATGCATATGTGTCTCCATACAAGTGGCTCATTAAAATAAGAACTTTGTAAACTGACTTAAAATCAGATAT
TTTTTCAAGAGTTAGGGAAAGTTGAAGTGTTTTACTGTTTTGTCTCTTGAGCCCTTTCTCTGGGGAAAAAATACA
TATCCATCTATCTATCTATATATAAACTGTGTATACATTCTTACTGTTTGAACAACTATTGCCTTTAATTAAATG
TTTCATTTTTCTCCAGAGTCCCCAAAGCCACATGGCATTATTATAGTCATTTTTGAGATGCCTGTAGAGAATGAA
AGTATTGACTCCGTTAGAGGGAAAATGGGTTTCTCTGGGTGAATTCCAACGAAGCATACCTAGGGGTAACAGTGA
ACCTACCTGGGTTTGTTTTGTTTTGGTAAGGATTTATGTAGTGTCTGGCTGTAAGCAAGAATGAGTGGATTATAA
ACTTGAAGATTTCTCTGTTAAAGTCACAAAAATGATCGACAAACAATATTTTTGTGATGTTTATTTAAACGTTGT
ATTTTATAACATACTTCAAGGAAGAGTATCGAAGTAAGTTGCTTTATAAATTAAGACTAAATTCGTATGGATGCA
GAATTCAATTAATAAAATTTGAGCCTGTTACGTAAATTGAATATTAATAAAATTGAAAATTTCAAAA
```

FIGURE 236

MENLSLSIEDVQPRSPGRSSLDDSGERDEKLSKSISFTSESISRVSETESFDGNSSKGGLGKE
ESQNEKQTKKSLLPTLEKKLTRVPSKSLDLNKNEYLSLDKSSTSDSVDEENVPEKDLHGRLFI
NRIFHISADRMFELLFTSSRFMQKFASSRNIIDVVSTPWTAELGGDQLRTMTYTIVLNSPLTG
KCTAATEKQTLYKESREARFYLVDSEVLTHDVPYHDYFYTVNRYCIIRSSKQKCRLRVSTDLK
YRKQPWGLVKSLIEKNSWSSLEDYFKQLESDLLIEESVLNQAIEDPGKLTGLRRRRRTFNRTA
ETVPKLSSQHSSGDVGLGAKGDITGKKKEMENYNVTLIVVMSIFVLLLVLLNVTLFLKLSKIE
HAAQSFYRLRLQEEKSLNLASDMVSRAETIQKNKDQAHRLKGVLRDSIVMLEQLKSSLIMLQK
TFDLLNKNKTGMAVES

Transmembrane domain:

amino acids 352-371

N-glycosylation sites.

amino acids 3-7, 54-58, 312-316, 349-353, 367-371, 449-453

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 81-85, 307-311

Tyrosine kinase phosphorylation sites.

amino acids 202-211, 246-254, 341-349

N-myristoylation site.

amino acids 259-265

Amidation site.

amino acids 339-343

FIGURE 237

CAGGGGCTGGAGGGCAGGGGAGGGG**ATG**ATGTCATTCCTGCTCGGCGCAATCCTGACCCTGCT
CTGGGCGCCCACGGCTCAGGCTGAGGTTCTGCTGCAGCCTGACTTCAATGCTGAAAAGTTCTC
AGGCCTCTGGTACGTGGTCTCCATGGCATCTGACTGCAGGGTCTTCCTGGGCAAGAAGGACCA
CCTGTCCATGTCCACCAGGGCCATCAGGCCCACAGAGGAGGGCGGCCTCCACGTCCACATGGA
GTTCCCGGGGGCGGACGGCTGTAACCAGGTGGATGCCGAGTACCTGAAGGTGGGCTCCGAGGG
ACACTTCAGAGTCCCGGCCTTGGGCTACCTGGACGTGCGCATCGTGGACACAGACTACAGCTC
CTTCGCCGTCCTTTACATCTACAAGGAGCTGGAGGGGGCCCTCAGCACCATGGTGCAGCTCTA
CAGCCGGACCCAGGATGTGAGTCCCCAGGCTCTGAAGTCCTTCCAGGACTTCTACCCGACCCT
GGGGCTCCCCAAGGACATGATGGTCATGCTGCCCCAGTCAGATGCATGCAACCCTGAGAGCAA
GGAGGCGCCC**TGA**CACCTCCGGAGCCCCACCCCCGCCCTTCCCAGGTGGAGCCAAAGCAGCAG
GCGCCTTTGCCCCTGGAGTCAAGACCCACAGCCCTCGGGGACCACCTGGAGTCTCTCCATCCT
CCACCCCCGCCTGTGGGATGCCTTGTGGGACGTCTCTTTCTATTCAATAAACAGATGCTGCA
GCCTCA

FIGURE 238

MMSFLLGAILTLLWAPTAQAEVLLQPDFNAEKFSGLWYVVSMASDCRVFLGKKDHLSMSTRAI
RPTEEGGLHVHMEFPGADGCNQVDAEYLKVGSEGHFRVPALGYLDVRIVDTDYSSFAVLYIYK
ELEGALSTMVQLYSRTQDVSPQALKSFQDFYPTLGLPKDMMVMLPQSDACNPESKEAP

Signal peptide:

amino acids 1-20

Tyrosine kinase phosphorylation site.

amino acids 110-117

N-myristoylation sites.

amino acids 7-13, 79-85, 130-136

Amidation site.

amino acids 50-54

FIGURE 239

GGCGCGCTGGTCCAGGTGAGCGGGCGCGTCCCCGCGACGGCGCTGCCTGCCCGAGGCGGTTCA

CGTAAAGACAGCGAGATCCTGAGGGCCAGCCGGGAAGGAGGCGTGGATATGGAGCTGGCTGCT

GCCAAGTCCGGGGCCCGCGCCGCTGCCTAGCGCGTCCTGGGGACTCTGTGGGGACGCGCCCCG

CGCCGCGGCTCGGGGACCCGTAGAGCCCGGCGCTGCGCGC**ATG**GCCCTGCTCTCGCGCCCCGC

GCTCACCCTCCTGCTCCTCCTCATGGCCGCTGTTGTCAGGTGCCAGGAGCAGGCCCAGACCAC

CGACTGGAGAGCCACCCTGAAGACCATCCGGAACGGCGTTCATAAGATAGACACGTACCTGAA

CGCCGCCTTGGACCTCCTGGGAGGCGAGGACGGTCTCTGCCAGTATAAATGCAGTGACGGATC

TAAGCCTTTCCCACGTTATGGTTATAAACCCTCCCCACCGAATGGATGTGGCTCTCCACTGTT

TGGTGTTCATCTTAACATTGGTATCCCTTCCCTGACAAAGTGTTGCAACCAACACGACAGGTG

CTATGAGACCTGTGGCAAAAGCAAGAATGACTGTGATGAAGAATTCCAGTATTGCCTCTCCAA

GATCTGCCGAGATGTACAGAAAACACTAGGACTAACTCAGCATGTTCAGGCATGTGAAACAAC

AGTGGAGCTCTTGTTTGACAGTGTTATACATTTAGGTTGTAAACCATATCTGGACAGCCAACG

AGCCGCATGCAGGTGTCATTATGAAGAAAAAACTGATCTT**TAA**AGGAGATGCCGACAGCTAGT

GACAGATGAAGATGGAAGAACATAACCTTTGACAAATAACTAATGTTTTTACAACATAAAACT

GTCTTATTTTTGTGAAAGGATTATTTTGAGACCTTAAAATAATTTATATCTTGATGTTAAAAC

CTCAAAGCAAAAAAAGTGAGGGAGATAGTGAGGGGAGGGCACGCTTGTCTTCTCAGGTATCTT

CCCCAGCATTGCTCCCTTACTTAGTATGCCAAATGTCTTGACCAATATCAAAAACAAGTGCTT

GTTTAGCGGAGAATTTTGAAAAGAGGAATATATAACTCAATTTTCACAACCACATTTACCAAA

AAAAGAGATCAAATATAAAATTCATCATAATGTCTGTTCAACATTATCTTATTTGGAAAATGG

GGAAATTATCACTTACAAGTATTTGTTTACTATGAAATTTTAAATACACATTTATGCCTAGAA

GGAACGGACTTTTTTTTTCTATTTTAATTACACATAATATGTAATTAAAGTACAACATAATAT

GTTGTTTCTCTGTAGCCCGTTGAGCATATGAGTAAGTCACATTTCTATTAGGACTACTTACAA

GGACAAGGTTTCCATTTTTCCAGTTGTAAAATTGGAACCATCAGCTGATAACCTCGTAGGGAG

CAACCCCAGGATAGCTAAGTGTTATGTAATATGCCTAGAAGGTGATGTGAATGCGATTCAGAA

GCATAGCCACTCCCATTTTATGAGCTACTCACATGACAAATGTCATCTTTTGCTATAACCTTT

GCCAAGTTAGAGAAAGATGGATTTAATGAGATAAATGAAAGATATTTAACCTAAAAAAAAA

AAAAAAAAAAAAAAAAA

FIGURE 240

MALLSRPALTLLLLLMAAVVRCQEQAQTTDWRATLKTIRNGVHKIDTYLNAALDLLGGEDGLC
QYKCSDGSKPFPRYGYKPSPPNGCGSPLFGVHLNIGIPSLTKCCNQHDRCYETCGKSKNDCDE
EFQYCLSKICRDVQKTLGLTQHVQACETTVELLFDSVIHLGCKPYLDSQRAACRCHYEEKTDL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites:

amino acids 57-63,93-99

Phospholipase A2 histidine active site:

amino acids 106-114

Neuraxin and MAP1B proteins repeat proteins Block:

amino acids 109-137

FIGURE 241

GATTCCGAGCGCCTCCACTGCTGGTCCGTTGGCCAGATCAACTCGCCGCGTGGGCCGGCCGTT

CCCTGAGAGTCTGAGCGCTCGCCGCACCCCCTTCCGAGCTTCTATTGGCCGTAGCAGACGTCC

GTCTGCCGCTATCTCCGCCCCAATACGGAAGCGGCCTAGTCCTCCGGCTCCGACAGCTGGGTG

TCCAGGCC**ATG**GGGCAGCCCTGGGCGGCTGGGAGCACGGACGGGGCGCCCGCGCAGCTGCCTC

TCGTGCTCACCGCGCTGTGGGCCGCGGCCGTGGGCCTGGAGCTGGCTTACGTGCTGGTGCTCG

GTCCCGGGCCGCCGCCGCTGGGACCCCTGGCCCGGGCCTTGCAGCTGGCGCTGGCCGCCTTCC

AGCTGCTCAACCTGCTGGGCAACGTGGGGCTCTTCCTGCGCTCGGATCCCAGCATCCGTGGCG

TGATGCTGGCCGGCCGCGGTCTGGGCCAGGGCTGGGCTTACTGCTACCAATGCCAAAGCCAGG

TGCCGCCACGCAGCGGACACTGCTCTGCCTGCCGCGTCTGCATCCTGCGTCGGGACCACCACT

GCCGCCTGCTGGGCCGCTGCGTGGGCTTCGGCAACTACCGGCCCTTCCTGTGCCTGCTGCTTC

ATGCCGCCGGCGTCCTGCTCCACGTCTCTGTGCTGCTGGGCCCTGCACTGTCGGCCCTGCTGC

GAGCCCACACGCCCCTCCACATGGCTGCCCTCCTCCTGCTTCCCTGGCTCATGTTGCTCACAG

GCAGAGTGTCTCTGGCACAGTTTGCCTTGGCCTTCGTGACGGACACGTGCGTGGCGGGTGCGC

TGCTGTGCGGGGCTGGGCTGCTCTTCCATGGGATGCTGCTGCTGCGGGGCCAGACCACATGGG

AGTGGGCTCGGGGCCAGCACTCCTATGACCTGGGTCCCTGCCACAACCTGCAGGCAGCCCTGG

GGCCCCGCTGGGCCCTCGTCTGGCTCTGGCCCTTCCTGGCCTCCCCATTGCCTGGGGATGGGA

TCACCTTCCAGACCACAGCAGATGTGGGACACACAGCCTCC**TGA**CTCCAGGAAGAGCCAGAGC

TGTGCAGGGAGGAAGGGGTGAGAGGGGGGCCCCCACACCTAGACTCAGTAAGGAAGTCGGGTT

GGACCTTAACATCTGCATTGGACAACTCCACCCCTTCCTTGGCCTTGCCCCTGCCCGCCTACA

CTCCTACGTGTCCAGGGCTTGGGCCGTGACTTAGGCAGAGGAGTGCAGAGGAGGGTCTGGCAG

GGGCTGCTCAGGCCGCCTAGCTGCCCCTTTGCCAGGTTAATAAAGCACTGACTTGTTAA

FIGURE 242

MGQPWAAGSTDGAPAQLPLVLTALWAAAVGLELAYVLVLGPGPPPLGPLARALQLALAAFQLL
NLLGNVGLFLRSDPSIRGVMLAGRGLGQGWAYCYQCQSQVPPRSGHCSACRVCILRRDHHCRL
LGRCVGFGNYRPFLCLLLHAAGVLLHVSVLLGPALSALLRAHTPLHMAALLLLPWLMLLTGRV
SLAQFALAFVTDTCVAGALLCGAGLLFHGMLLLRGQTTWEWARGQHSYDLGPCHNLQAALGPR
WALVWLWPFLASPLPGDGITFQTTADVGHTAS

Important features:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 51-66,143-160,174-191,198-214

N-myristoylation sites:

amino acids 2-8,8-14,30-36,81-87,88-94,90-96,206-212

Leucine zipper pattern:

amino acids 143-165,150-172,157-179,164-186

FIGURE 243

CTTGTCTTTGTGTCGGTTGTGATTTTCCTAATCTCTGATTTTCCTTTTCTCTCGGACGCTCTC

CCTCTTCGGACCCATTTTCTCCCGTGCTTCATGCCCTGATAGCCTGGCCCCTTCCCGGCTTCC

TTCGCTACCGGGGACGCCTCTAGTTTTTCTGAATTTCTGGCTGGCTCCACCCTCCGCGTTCAT

CTTCCTCAAGAGTTCGCCCCTCTGGGGGCTCCTCTGTGTAATCGTCGCCTTCTCTGGGTATTT

CTGTGAACTCCGTCTCACACCATCCCGCCATCTTCTCTGCCTTGGCCCCTTTTCTCTGTACAG

CCAGCTCTGTGTCCTTTTCTTCTCCCCCTCTAAAATCGACTCCTCTTCTCCCTGAGAGCCCCA

CCTTTGTGCCCCACTCCTCATTTTCCTACGCCTCCCTCTCTCTGCTGGTCCTCTCTCTCCCTG

CAAGGTTCCATTCCATCAATTTGTTTGTCTTTTGTAGGGGTGGCATCCCCTCTGACTACTGCT

CCATCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCTTGAGGATTTCACTTCAATCTTTTCTGGT

TGCGTCTCCACTTGTACTCAGCTTGTTAGGTCCAGGTCCAGTTGTTCTGCATCTGAGGCTGGC

GTGTGCTGTCTTCTCTGATTGGCCTAATCTCCCTCACCCCCGTGAGATCTGTTGTCAGCCTTC

GTTTCTCTTTCCTGTGTCCCAGCTTTTCTGCGGGTCTTGGCACCTTTCTTGGCCACAGATTTC

TGGGTTACAGAGCATGTGTGTCTGAGGCATTGCAGGCAGAAAAGGGTGGCCGACGTGACCTCT

AGCTGGACTGCTGGGCAGGGGAGCTGTCCTAGATAAAATTGGAAAGAAACAGTGACCCAGAGA

CAGGTGGACAAAGAATTCGGGGACTGATGGGAACTGAGCTTGGGATCCAGACTGAAACTGATT

CCAGACTGACCTCTAGCACCCAGGACCCAGACACAGGGCC**ATG**GGACCCCAGCATTTGAGACT

TGTGCAGCTGTTCTGCCTTCTAGGGGCCATCCCCACTCTGCCTCGGGCTGGAGCTCTTTTGTG

CTATGAAGCAACAGCCTCAAGATTCAGAGCTGTTGCTTTCCATAACTGGAAGTGGCTTCTGAT

GAGGAACATGGTGTGTAAGCTGCAAGAGGGCTGCGAGGAGACGCTAGTGTTCATTGAGACAGG

GACTGCAAGGGGAGTTGTGGGCTTTAAAGGCTGCAGCTCGTCTTCGTCTTACCCTGCGCAAAT

CTCCTACCTTGTTTCCCCACCCGGAGTGTCCATTGCCTCCTACAGTCGCGTCTGCCGGTCTTA

TCTCTGCAACAACCTCACCAATTTGGAGCCTTTTGTGAAACTCAAGGCCAGCACTCCTAAGTC

TATCACATCTGCGTCCTGTAGCTGCCCGACCTGTGTGGGCGAGCACATGAAGGATTGCCTCCC

AAATTTTGTCACCACTAATTCTTGCCCCTTGGCTGCTTCTACGTGTTACAGTTCCACCTTAAA

ATTTCAGGCAGGGTTTCTCAATACCACCTTCCTCCTCATGGGGTGTGCTCGTGAACATAACCA

GCTTTTAGCAGATTTTCATCATATTGGGAGCATCAAAGTGACTGAGGTCCTCAACATCTTAGA

GAAGTCTCAGATTGTTGGTGCAGCATCCTCCAGGCAAGATCCTGCTTGGGGTGTCGTCTTAGG

CCTCCTGTTTGCCTTCAGGGAC**TGA**CCATCTAGCTGCACCCGACAAGCACCCAGACTCTTTCA

CATAACAAATAAAATAGCAGAGTTCCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAA

FIGURE 244

MGPQHLRLVQLFCLLGAIPTLPRAGALLCYEATASRFRAVAFHNWKWLLMRNMVCKLQEGCEE

TLVFIETGTARGVVGFKGCSSSSSYPAQISYLVSPPGVSIASYSRVCRSYLCNNLTNLEPFVK

LKASTPKSITSASCSCPTCVGEHMKDCLPNFVTTNSCPLAASTCYSSTLKFQAGFLNTTFLLM

GCAREHNQLLADFHHIGSIKVTEVLNILEKSQIVGAASSRQDPAWGVVLGLLFAFRD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation sites:

amino acids 117-121,183-187

N-myristoylation sites:

amino acids 16-22,25-31,60-66,71-77,81-87,100-106,224-230,
235-241,239-245

Prokaryotic membrane lipoprotein lipid attachment site:

amino acids 181-192

FIGURE 245

GTGGAGTTGGGTGGTGTCGGGAGCCTCTCCCTGAGGGGCACCGCGTCTTCAGGAGCTGGGCCTCCAGTGCGGCGC
GATGTCAGGCGCGGTGACAGCTCTGTGAGTCCGAGGCCGCGGCCGTGGCGCTGGGCGGCTGCGGGGCCTGACCGG
TCCGCTC**ATG**GTGCCGCCACGACGCCATCGCGGGGCAGGAAGGCCAGGGGTGCTGAGTTCTTCACCTCCTTTTAG
ACTGAGATCTGCCAAGTTTTCCGGCATTGCTCTTGAGGATCTCAGAAGGGCTCTTAAGACAAGACTGCAAATGGT
GTGTGTATTTGTCATGAACCGAATGAATTCCCAGAACAGTGGTTTCACTCAGCGCAGGCGAATGGCTCTTGGGAT
TGTTATTCTTCTGCTTGTTGATGTGATATGGGTTGCTTCCTCTGAACTTACTTCGTATGTTTTTACCCAGTACAA
CAAACCATTCTTCAGCACCTTTGCAAAAACATCTATGTTTGTTTTGTACCTTTTGGGCTTTATTATTTGGAAGCC
ATGGAGACAACAGTGTACAAGAGGACTTCGCGGAAAGCATGCTGCTTTTTTTGCAGATGCTGAAGGTTACTTTGC
TGCTTGCACAACAGATACAACTATGAATAGTTCTTTGAGTGAACCTCTGTATGTGCCTGTGAAATTCCATGATCT
TCCAAGTGAAAAACCTGAGAGCACAAACATTGATACTGAAAAAACCCCCAAAAAGTCTCGTGTGAGGTTCAGTAA
TATCATGGAGATTCGACAGCTTCCGTCAAGTCATGCATTGGAAGCAAAGTTGTCTCGCATGTCATATCCTGTGAA
AGAACAAGAATCCATACTGAAAACTGTGGGGAAACTTACTGCAACTCAAGTAGCGAAAATTAGCTTTTTTTTTTG
CTTTGTGTGGTTTTTGGCAAATTTGTCATATCAAGAAGCACTTTCAGACACACAAGTTGCTATAGTTAATATTTT
ATCTTCAACTTCCGGACTTTTTACCTTAATCCTTGCTGCAGTATTTCCAAGTAACAGTGGAGATAGATTTACCCT
TTCTAAACTATTAGCTGTAATTTTAAGCATTGGAGGCGTTGTACTGGTAAACCTGGCAGGGTCTGAAAAACCTGC
TGGAAGAGACACAGTAGGTTCCATTTGGTCTCTTGCTGGAGCCATGCTCTATGCTGTCTATATTGTTATGATTAA
GAGAAAAGTAGATAGAGAAGACAAGTTGGATATTCCAATGTTCTTTGGTTTTGTAGGTTTGTTTAATCTGCTGCT
CTTATGGCCAGGTTTCTTTTTACTTCATTATACTGGATTTGAGGACTTCGAGTTTCCCAATAAAGTAGTATTAAT
GTGCATTATCATTAATGGCCTTATTGGAACAGTACTCTCAGAGTTCCTGTGGTTGTGGGGCTGCTTTCTTACCTC
ATCATTGATAGGCACACTTGCACTAAGCCTTACAATACCTCTGTCCATAATAGCTGACATGTGTATGCAAAAGGT
GCAGTTTTCTTGGTTATTTTTTGCAGGAGCTATCCCTGTATTTTTTTCATTTTTTATTGTAACTCTCCTATGCCA
TTATAATAATTGGGATCCTGTGATGGTGGGAATCAGAAGAATATTTGCTTTTATATGCAGAAAACATCGAATTCA
GAGAGTTCCAGAAGACAGCGAACAGTGTGAGAGTCTCATTTCTATGCACAGTGTTTCTCAGGAGGATGGAGCTAG
T**TAG**CTGTCTGTTGTCTGTAGCCCAGCTTGATAATGGAACTATACAGCGAAGAGACAATCTCTGGCAAGTTTTTG
TAGAAAAAATGTTTCAGTGCCTAGTCTGAAAAATAACAGTTTGAGTTCTTTGAAACTCTAAAATATATTTTTCTC
ATACCTGTTTTCTTCATTTTCATAATGAAGCACTTTGCTATGTAGCTGTGTACATATCACTACAGTTATAGGAAG
TTTCAGTCTACAGTCCATCCAAAGGACCAACCTGCCTTACACATCTCAAGGAATTCAGCTGTTGAAATCATTTGA
ACTAATCAAGGAATAAATCCTAATGTTCTGGGACTTTATTTTCACATGTTAAATGCTGGAATATATTATGAAAAT
GTTTTCAAGAAATCACTTAAGTGTTCATAGACCAGTATTTCTGACAGGTAAAATGCTAAAATAAGCTACCTGTAA
TAAGTGTGGATTATATTTTTGGGTTTTGTAGAATATTGCAAATTAACCACACAAAAAATGTTTAATTTATGCAAC
AAGCATGTTTGTGCAAATTTCATGGGACTTTAAAAAGAATAAGTATTTGAGAAAATATCTGGTTCACTTACACTA
CATTTACTGTATTATTCTTTTATAGCATTAGGTGCCTTGTATTTTAAATCTGTGACAAACCATGGCAAATTTTTA
AAGGGGAAGTATTATTATAAAATGAAGAAATATGTATTTCTAAAGGCTATATTGCTGTAAACTTAATTGATAAAG
CTCTGTTTAATTTAGAGTTTTGAAGAAATAGTCTCCCTTCAATTAAGAAATTTTCATAATGGAATGATTTAAATT
GAAGTGACAAAGAGTATTATTAAAATACAATGTTTATAAAAAAA

FIGURE 246

MVPPRRHRGAGRPGVLSSSPPFRLRSAKFSGIALEDLRRALKTRLQMVCVFVMNRMNSQNSGF
TQRRRMALGIVILLLVDVIWVASSELTSYVFTQYNKPFFSTFAKTSMFVLYLLGFIIWKPWRQ
QCTRGLRGKHAAFFADAEGYFAACTTDTTMNSSLSEPLYVPVKFHDLPSEKPESTNIDTEKTP
KKSRVRFSNIMEIRQLPSSHALEAKLSRMSYPVKEQESILKTVGKLTATQVAKISFFFCFVWF
LANLSYQEALSDTQVAIVNILSSTSGLFTLILAAVFPSNSGDRFTLSKLLAVILSIGGVVLVN
LAGSEKPAGRDTVGSIWSLAGAMLYAVYIVMIKRKVDREDKLDIPMFFGFVGLFNLLLLWPGF
FLLHYTGFEDFEFPNKVVLMCIIINGLIGTVLSEFLWLWGCFLTSSLIGTLALSLTIPLSIIA
DMCMQKVQFSWLFFAGAIPVFFSFFIVTLLCHYNNWDPVMVGIRRIFAFICRKHRIQRVPEDS
EQCESLISMHSVSQEDGAS

Important features:

Transmembrane domain:

amino acids 69-87,105-118,237-256,266-285,300-316,332-346,
364-379,399-419,453-472

N-glycosylation sites:

amino acids 157-161,255-259

N-myristoylation sites:

amino acids 14-20,329-335,404-410,407-413,418-424

FIGURE 247

CGTCTGTAGAGATATCATGAACTTCAACTTAGCTTTGGTACTTTCTTCCCTGAAGACAGAGGG

CAGAACTCTGAGTTCCAGAACCATTTTCAACTGTATTGGGGACCAATCACTTGACTCTATTCT

TGTCTCTCTGACAGATGACGCTACACTCTCCTCTGAATAATGGACACCATTTCTAAAACTGAA

TCCTGCTACTAAAATAATTCAGATGATATATTTTTCCAATTCTACAATCTTGCTTTGTTTTAT

TTAGTTGTTTTCTCTCTCTCTTCCCAGTTTTCCAGAGACTGGAGCTAAACTGGGCTTTCAACA

TCATC**ATG**AAGTTTATCCTCCTCTGGGCCCTCTTGAATCTGACTGTTGCTTTGGCCTTTAATC

CAGATTACACAGTCAGCTCCACTCCCCCTTACTTGGTCTATTTGAAATCTGACTACTTGCCCT

GCGCTGGAGTCCTGATCCACCCGCTTTGGGTGATCACAGCTGCACACTGCAATTTACCAAAGC

TTCGGGTGATATTGGGGGTTACAATCCAGCAGACTCTAATGAAAAGCATCTGCAAGTGATTG

GCTATGAGAAGATGATTCATCATCCACACTTCTCAGTCACTTCTATTGATCATGACATCATGC

TAATCAAGCTGAAAACAGAGGCTGAACTCAATGACTATGTGAAATTAGCCAACCTGCCCTACC

AAACTATCTCTGAAAATACCATGTGCTCTGTCTCTACCTGGAGCTACAATGTGTGTGATATCT

ACAAAGAGCCCGATTCACTGCAAACTGTGAACATCTCTGTAATCTCCAAGCCTCAGTGTCGCG

ATGCCTATAAAACCTACAACATCACGGAAAATATGCTGTGTGTGGGCATTGTGCCAGGAAGGA

GGCAGCCCTGCAAGGAAGTTTCTGCTGCCCCGGCAATCTGCAATGGGATGCTTCAAGGAATCC

TGTCTTTTGCGGATGGATGTGTTTTGAGAGCCGATGTTGGCATCTATGCCAAAATTTTTTACT

ATATACCCTGGATTGAAAATGTAATCCAAAATAAC**TGA**GCTGTGGCAGTTGTGGACCATATGA

CACAGCTTGTCCCCATCGTTCACCTTTAGAATTAAATATAAATTAACTCCTC

FIGURE 248

MKFILLWALLNLTVALAFNPDYTVSSTPPYLVYLKSDYLPCAGVLIHPLWVITAAHCNLPKLR
VILGVTIPADSNEKHLQVIGYEKMIHHPHFSVTSIDHDIMLIKLKTEAELNDYVKLANLPYQT
ISENTMCSVSTWSYNVCDIYKEPDSLQTVNISVISKPQCRDAYKTYNITENMLCVGIVPGRRQ
PCKEVSAAPAICNGMLQGILSFADGCVLRADVGIYAKIFYYIPWIENVIQNN

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation sites:

amino acids 11-15,156-160,173-177

Tyrosine kinase phosphorylation site:

amino acids 108-117

N-myristoylation sites:

amino acids 182-188,203-209

Amidation site:

amino acids 185-189

Serine proteases, trypsin family, histidine active site:

amino acids 52-58

FIGURE 249

GCGAGGCGGCCGCTGTCTTCTGCTGCGGCTTCCGCGACCACAAGTACTGCTGCGACGACCCGC

ACAGCTTCTTCCCCTACGAGCACAGCTAC**ATG**TGGTGGCTCAGCATTGGCGCTCTCATAGGCC

TGTCCGTAGCAGCAGTGGTTCTTCTCGCCTTCATTGTTACCGCCTGTGTGCTCTGCTACCTGT

TCATCAGCTCTAAGCCCCACACAAAGTTGGACCTGGGCTTGAGCTTACAGACAGCAGGCCCTG

AGGAGGTTTCTCCTGACTGCCAAGGTGTGAACACAGGCATGGCGGCAGAAGTGCCAAAAGTGA

GCCCTCTCCAGCAGAGTTACTCCTGCTTGAACCCGCAGCTGGAGAGCAATGAGGGGCAGGCTG

TGAACTCCAAACGCCTCCTCCATCATTGCTTCATGGCCACAGTGACCACCAGTGACATTCCAG

GCAGCCCTGAGGAAGCCTCTGTACCCAACCCTGACCTATGTGGACCAGTCCCA**TAA**ACATTCA

ATAAATGTCTCCATACCATCAA

FIGURE 250

MWWLSIGALIGLSVAAVVLLAFIVTACVLCYLFISSKPHTKLDLGLSLQTAGPEEVSPDCQGV
NTGMAAEVPKVSPLQQSYSCLNPQLESNEGQAVNSKRLLHHCFMATVTTSDIPGSPEEASVPN
PDLCGPVP

Important features:

Signal peptide:

Amino acids 1-26

N-myristoylation sites:

Amino acids 7-13,11-17,62-68,93-99

FIGURE 251

GTGGTTTGGATTGAGCCGGGCCCGGCCGGGGCGCCGAGTCGGAGGGGGTGGCAGTGAGCGGCG

GCAGAGGCTACGGGGCTCGGTTTGGCTGACTGGGGAGTCGGCAGGCGGCAGGAACC**ATG**CGAG

GCCAGCGGAGCCTGCTGCTGGGCCCGGCCCGCCTCTGCCTCCGCCTCCTTCTGCTGCTGGGTT

ACAGGCGCCGCTGTCCACCTCTACTCCGGGGTCTAGTACAGCGCTGGCGCTACGGCAAGGTCT

GCCTGCGCTCCCTGCTCTACAACTCCTTTGGGGGCAGTGACACCGCTGTTGATGCTGCCTTTG

AGCCTGTCTACTGGCTGGTAGACAACGTGATCCGCTGGTTTGGAGTGGTGTTCGTGGTCCTGG

TGATCGTGCTGACAGGCTCCATTGTAGCTATCGCCTACCTGTGTGTCCTGCCTCTCATCCTCC

GAACCTACTCAGTGCCACGACTCTGCTGGCATTTCTTCTATAGCCACTGGAATCTGATCCTGA

TTGTCTTCCACTACTACCAGGCCATCACCACTCCGCCTGGGTACCCACCCCAGGGCAGGAATG

ATATCGCCACCGTCTCCATCTGTAAGAAGTGCATTTACCCCAAGCCAGCCCGAACACACCACT

GCAGCATCTGCAACAGGTGTGTGCTGAAGATGGATCACCACTGCCCCTGGCTAAACAATTGTG

TGGGCCACTATAACCATCGGTACTTCTTCTCTTTCTGCTTTTTCATGACTCTGGGCTGTGTCT

ACTGCAGCTATGGAAGTTGGGACCTTTTCCGGGAGGCTTATGCTGCCATTGAGACTTATCACC

AGACCCCACCACCCACCTTCTCCTTTCGAGAAAGGATGACTCACAAGAGTCTTGTCTACCTCT

GGTTCCTGTGCAGTTCTGTGGCACTTGCCCTGGGTGCCCTAACTGTATGGCATGCTGTTCTCA

TCAGTCGAGGTGAGACTAGCATCGAAAGGCACATCAACAAGAAGGAGAGACGTCGGCTACAGG

CCAAGGGCAGAGTATTTAGGAATCCTTACAACTACGGCTGCTTGGACAACTGGAAGGTATTCC

TGGGTGTGGATACAGGAAGGCACTGGCTTACTCGGGTGCTCTTACCTTCTAGTCACTTGCCCC

ATGGGAATGGAATGAGCTGGGAGCCCCCTCCCTGGGTGACTGCTCACTCAGCCTCTGTGATGG

CAGTG**TGA**GCTGGACTGTGTCAGCCACGACTCGAGCACTCATTCTGCTCCCTATGTTATTTCA

AGGGCCTCCAAGGGCAGCTTTTCTCAGAATCCTTGATCAAAAAGAGCCAGTGGGCCTGCCTTA

GGGTACCATGCAGGACAATTCAAGGACCAGCCTTTTTACCACTGCAGAAGAAAGACACAATGT

GGAGAAATCTTAGGACTGACATCCCTTTACTCAGGCAAACAGAAGTTCCAACCCCAGACTAGG

GGTCAGGCAGCTAGCTACCTACCTTGCCCAGTGCTGACCCGGACCTCCTCCAGGATACAGCAC

TGGAGTTGGCCACCACCTCTTCTACTTGCTGTCTGAAAAAACACCTGACTAGTACAGCTGAGA

TCTTGGCTTCTCAACAGGGCAAAGATACCAGGCCTGCTGCTGAGGTCACTGCCACTTCTCACA

TGCTGCTTAAGGGAGCACAAATAAAGGTATTCGATTTTTAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAA

FIGURE 252

MRGQRSLLLGPARLCLRLLLLLGYRRRCPPLLRGLVQRWRYGKVCLRSLLYNSFGGSDTAVDA
AFEPVYWLVDNVIRWFGVVFVVLVIVLTGSIVAIAYLCVLPLILRTYSVPRLCWHFFYSHWNL
ILIVFHYYQAITTPPGYPPQGRNDIATVSICKKCIYPKPARTHHCSICNRCVLKMDHHCPWLN
NCVGHYNHRYFFSFCFFMTLGCVYCSYGSWDLFREAYAAIETYHQTPPPTFSFRERMTHKSLV
YLWFLCSSVALALGALTVWHAVLISRGETSIERHINKKERRRLQAKGRVFRNPYNYGCLDNWK
VFLGVDTGRHWLTRVLLPSSHLPHGNGMSWEPPPWVTAHSASVMAV

Important features:

Transmembrane domain:

amino acids 88-100,202-216,254-274

N-myristoylation sites:

amino acids 55-61,56-62,92-98,210-216,309-315,319-325,340-346

Prokaryotic membrane lipoprotein lipid attachment site:

amino acids 201-212

FIGURE 253

GATCAAGCGCCTTCCTTTCCCTTCCTCTCCCTACTTGGCCTTTGCCCTAAGCCAAGACCTGGCCATCAGCCTGGC

TGCAGGGGCCTGCAGAGCCAGCTGCACTTTTTCAGGTATGGGGGAGGGCCAGGCACC**ATG**AAGCCAGTGTGGGTC

GCCACCCTTCTGTGGATGCTACTGCTGGTGCCCAGGCTGGGGGCCGCCCGGAAGGGGTCCCCAGAAGAGGCCTCC

TTCTACTATGGAACCTTCCCTCTTGGCTTCTCCTGGGGCGTGGGCAGTTCTGCCTACCAGACGGAGGGCGCCTGG

GACCAGGACGGGAAAGGGCCTAGCATCTGGGACGTCTTCACACACAGTGGGAAGGGGAAAGTGCTTGGGAATGAG

ACGGCAGATGTAGCCTGTGACGGCTACTACAAGGTCCAGGAGGACATCATTCTGCTGAGGGAACTGCACGTCAAC

CACTACCGATTCTCCCTGTCTTGGCCCCGGCTCCTGCCCACAGGCATCCGAGCCGAGCAGGTGAACAAGAAGGGA

ATCGAATTCTACAGTGATCTTATCGATGCCCTTCTGAGCAGCAACATCACTCCCATCGTGACCTTGCACCACTGG

GATCTGCCACAGCTGCTCCAGGTCAAATACGGTGGGTGGCAGAATGTGAGCATGGCCAACTACTTCAGAGACTAC

GCCAACCTGTGCTTTGAGGCCTTTGGGGACCGTGTGAAGCACTGGATCACGTTCAGTGATCCTCGGGCAATGGCA

GAAAAAGGCTATGAGACGGGCCACCATGCGCCGGGCCTGAAGCTCCGCGGCACCGGCCTGTACAAGGCAGCACAC

CACATCATTAAGGCCCACGCCAAAACCTGGCATTCTTATAACACCACGTGGCGCAGCAAGCAGCAAGGTCTGGTG

GGAATTTCACTGAACTGTGACTGGGGGGGAACCTGTGGACATTAGTAACCCCAAGGACCTAGAGGCTGCCGAGAGA

TACCTACAGTTCTGTCTGGGCTGGTTTGCCAACCCCATTTATGCCGGTGACTACCCCCAAGTCATGAAGGACTAC

ATTGGAAGAAAGAGTGCAGAGCAAGGCCTGGAGATGTCGAGGTTACCGGTGTTCTCACTCCAGGAGAAGAGCTAC

ATTAAAGGCACATCCGATTTCTTGGGATTAGGTCATTTTACTACTCGGTACATCACGGAAAGGAACTACCCCTCC

CGCCAGGGGCCCAGCTACCAGAACGATCGTGACTTGATAGAGCTGGTTGACCCAAACTGGCCAGATCTGGGGTCT

AAATGGCTATATTCTGTGCCATGGGGATTTAGGAGGCTCCTTAACTTTGCTCAGACTCAATACGGTGATCCTCCC

ATATATGTGATGGAAAATGGAGCATCTCAAAAATTCCACTGTACTCAATTATGTGATGAGTGGAGAATTCAATAC

CTTAAAGGATACATAAATGAAATGCTAAAAGCTATAAAAGATGGTGCTAATATAAAGGGGTATACTTCCTGGTCT

CTGTTGGATAAGTTTGAATGGGAGAAAGGATACTCAGATAGATATGGATTCTACTATGTTGAATTTAACGACAGA

AATAAGCCTCGCTATCCAAAGGCTTCAGTTCAATATTACAAGAAGATTATCATTGCCAATGGGTTTCCCAATCCA

AGAGAGGTGGAAAGTTGGTACCTCAAAGCTTTGGAAACTTGCTCTATCAACAATCAGATGCTTGCTGCAGAGCCT

TTGCTAAGTCACATGCAAATGGTTACGGAGATCGTGGTACCCACTGTCTGCTCCCTCTGTGTCCTCATCACTGCT

GTTCTACTAATGCTCCTCCTGAGGAGGCAGAGC**TGA**GACAGGATTATCAATTTTGGAGCTTCATAAGAGAATCTT

CAGGATCTTCCTCCCTTTTCTGCTTTGAGGGTTTCCATACATTGCTGTTTTCAGGTTCTACAATAATTACCTTTT

TTTCTCTTTCTCTTTTTGGCTTGTGCTGGGATTTAAGAATTAGAAAATAAAAATAAGCAGAAATTA

FIGURE 254

MKPVWVATLLWMLLLVPRLGAARKGSPEEASFYYGTFPLGFSWGVGSSAYQTEGAWDQDGKGPSIWDVFTHSGKG
KVLGNETADVACDGYYKVQEDIILLRELHVNHYRFSLSWPRLLPTGIRAEQVNKKGIEFYSDLIDALLSSNITPI
VTLHHWDLPQLLQVKYGGWQNVSMANYFRDYANLCFEAFGDRVKHWITFSDPRAMAEKGYETGHHAPGLKLRGTG
LYKAAHHIIKAHAKTWHSYNTTWRSKQQGLVGISLNCDWGEPVDISNPKDLEAAERYLQFCLGWFANPIYAGDYP
QVMKDYIGRKSAEQGLEMSRLPVFSLQEKSYIKGTSDFLGLGHFTTRYITERNYPSRQGPSYQNDRDLIELVDPN
WPDLGSKWLYSVPWGFRRLLNFAQTQYGDPPIYVMENGASQKFHCTQLCDEWRIQYLKGYINEMLKAIKDGANIK
GYTSWSLLDKFEWEKGYSDRYGFYYVEFNDRNKPRYPKASVQYYKKIIIANGFPNPREVESWYLKALETCSINNQ
MLAAEPLLSHMQMVTEIVVPTVCSLCVLITAVLLMLLLRRQS

Important features:
Signal peptide:
amino acids 1-21

Transmembrane domain:
amino acids 541-558

N-glycosylation sites:
amino acids 80-84,171-175,245-249

Glycosaminoglycan attachment site:
amino acids 72-76

cAMP- and cGMP-dependent protein kinase phosphorylation sites:
amino acids 23-27,564-568

Tyrosine kinase phosphorylation sites:
amino acids 203-211,347-355,460-468,507-514

N-myristoylation sites:
amino acids 44-50,79-85,167-173,225-231,257-263,315-321

Amidation site:
amino acids 307-311

Glycosyl hydrolases family 1 active site:
amino acids 407-416

Glycosyl hydrolases family 1 N-terminal signature:
amino acids 41-56

Motif name Glycosyl hydrolases family:
amino acids 37- 67

FIGURE 255

CGCGAAG**ATG**CGAAAGGTGGTTTTGATCACCGGGGCTAGCAGTGGCATTGGCCTGGCCCTCTG

CAAGCGGCTGCTGGCGGAAGATGATGAGCTTCATCTGTGTTTGGCGTGCAGGAACATGAGCAA

GGCAGAAGCTGTCTGTGCTGCTCTGCTGGCCTCTCACCCCACTGCTGAGGTCACCATTGTCCA

GGTGGATGTCAGCAACCTGCAGTCGGTCTTCCGGGCCTCCAAGGAACTTAAGCAAAGGTTTCA

GAGATTAGACTGTATATATCTAAATGCTGGGATCATGCCTAATCCACAACTAAATATCAAAGC

ACTTTTCTTTGGCCTCTTTTCAAGAAAAGTGATTCATATGTTCTCCACAGCTGAAGGCCTGCT

GACCCAGGGTGATAAGATCACTGCTGATGGACTTCAGGAGGTGTTTGAGACCAATGTCTTTGG

CCATTTTATCCTGATTCGGGAACTGGAGCCTCTCCTCTGTCACAGTGACAATCCATCTCAGCT

CATCTGGACATCATCTCGCAGTGCAAGGAAATCTAATTTCAGCCTCGAGGACTTCCAGCACAG

CAAAGGCAAGGAACCCTACAGCTCTTCCAAATATGCCACTGACCTTTTGAGTGTGGCTTTGAA

CAGGAACTTCAACCAGCAGGGTCTCTATTCCAATGTGGCCTGTCCAGGTACAGCATTGACCAA

TTTGACATATGGAATTCTGCCTCCGTTTATATGGACGCTGTTGATGCCGGCAATATTGCTACT

TCGCTTTTTTGCAAATGCATTCACTTTGACACCATATAATGGAACAGAAGCTCTGGTATGGCT

TTTCCACCAAAAGCCTGAATCTCTCAATCCTCTGATCAAATATCTGAGTGCCACCACTGGCTT

TGGAAGAAATTATATTATGACCCAGAAGATGGACCTAGATGAAGACACTGCTGAAAAATTTTA

TCAAAAGTTACTGGAACTGGAAAAGCACATTAGGGTCACTATTCAAAAAACAGATAATCAGGC

CAGGCTCAGTGGCTCATGCCTA**TAA**TTCCAGCACTTTGGGAGGCCAAGGCAGAAGGATCACTT

GAGACCAGGAGTTCAAGACCAGCCTGAGAACATAGTGAGCCCTTGTCTCTACAAAAAGAAAT

AAAAATAATAGCTGGGTGTGGTGGCATGCGCATGTAGTCCCAGCTACTCAGAAGGATGAGGTG

GGAGGATCTCTTGAGGCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCACTGCACTCC

AGCCTGGGTGACAGCGAGACCCTGTCTCAAAATATGTATATATTTAATATATATATAAAACCA

GAGCTGACAATGACACTCTGGAACATTGCATACCTTCTGTACATTCTGGGGTACATGGATTTC

TACTGAGTTGGATAATATGCATTTGTAATAAACTATGAACTATGAA

FIGURE 256

MRKVVLITGASSGIGLALCKRLLAEDDELHLCLACRNMSKAEAVCAALLASHPTAEVTIVQVD
VSNLQSVFRASKELKQRFQRLDCIYLNAGIMPNPQLNIKALFFGLFSRKVIHMFSTAEGLLTQ
GDKITADGLQEVFETNVFGHFILIRELEPLLCHSDNPSQLIWTSSRSARKSNFSLEDFQHSKG
KEPYSSSKYATDLLSVALNRNFNQQGLYSNVACPGTALTNLTYGILPPFIWTLLMPAILLLRF
FANAFTLTPYNGTEALVWLFHQKPESLNPLIKYLSATTGFGRNYIMTQKMDLDEDTAEKFYQK
LLELEKHIRVTIQKTDNQARLSGSCL

Important features:

Transmembrane domain:

amino acids 234-254

N-glycosylation sites:

amino acids 37-41,178-182,229-233,263-267

Glycosaminoglycan attachment site:

amino acids 12-16

N-myristoylation sites:

amino acids 9-15,13-19,15-21,215-221,224-230

FIGURE 257

CGGACGCGTGGGGCCGT**ATG**CGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCC

CCAGCCCCCTGCTCCTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGA

AGACCCGCCAGGTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTC

ATATACGGGCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGG

CAGTGGTAATGGTGGCCACCAACACCCCCACAGCACCCTGAGCATCAACTGGAGCCTCCTGC

TATCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCTG

CCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCAAAGC

CTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATCACTGATT

CATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGACCCTACCAGGA

CTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCCAGCCGACCAGCCC

AACCCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTGGCCCTGATTGGAGCCT

CTCCCCGGGGAAACCGTTCCCTGTTTGGGCTGGAGGTAGCCACATTGGGCCAGGGCCCTGACT

GCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATATGCACCGGCCGTCTTCCAGTTGG

ACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCACAGTGGCGACCAGTGGCTTACTCCC

AGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGCCAAGCTTCCCCTCTTCATCCTGCCTTAG

CATACTCTCTTCCCCAGTCACCCATTGTCCGAGCCTTCTTTGGGTCCCAGAATAACTTCTGTG

CCTTCAATCTGACGTTCGGGGCTTCCACAGGCCCTGGCTATTGGGACCAACACTACCTCAGCT

GGTCGATGCTCCTGGGTGTGGGCTTCCCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCA

TCATGGCAGTGGCCCTGGGTGCCCCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGC

TGCACCACAAGAAGTACTCAGAGTACCAGTCCATAAAT**TAA**GGCCCGCTCTCTGGAGGGAAGG

ACATTACTGAACCTGTCTTGCTGTGCCTCGAAACTCTGGAGGTTGGAGCATCAAGTTCCAGCC

GGCCCCTTCACTCCCCCATCTTGCTTTTCTGTGGAACCTCAGAGGCCAGCCTCGACTTCCTGG

AGACCCCCAGGTGGGGCTTCCTTCATACTTTGTTGGGGGACTTTGGAGGCGGGCAGGGGACAG

GGCTATTGATAAGGTCCCCTTGGTGTTGCCTTCTTGCATCTCCACACATTTCCCTTGGATGGG

ACTTGCAGGCCTAAATGAGAGGCATTCTGACTGGTTGGCTGCCCTGGAAGGCAAGAAAATAGA

TTTATTTTTTTTCACAGGGAAAAAAAAAAAA

FIGURE 258

MRGSVECTWGWGHCAPSPLLLWTLLLFAAPFGLLGEKTRQVSLEVIPNWLGPLQNLLHIRAVG
TNSTLHYVWSSLGPLAVVMVATNTPHSTLSINWSLLLSPEPDGGLMVLPKDSIQFSSALVFTR
LLEFDSTNVSDTAAKPLGRPYPPYSLADFSWNNITDSLDPATLSATFQGHPMNDPTRTFANGS
LAFRVQAFSRSSRPAQPPRLLHTADTCQLEVALIGASPRGNRSLFGLEVATLGQGPDCPSMQE
QHSIDDEYAPAVFQLDQLLWGSLPSGFAQWRPVAYSQKPGGRESALPCQASPLHPALAYSLPQ
SPIVRAFFGSQNNFCAFNLTFGASTGPGYWDQHYLSWSMLLGVGFPPVDGLSPLVLGIMAVAL
GAPGLMLLGGGLVLLLHHKKYSEYQSIN

Important features:

Signal peptide:

amino acids 1-35

Transmembrane domain:

amino acids 365-386

N-glycosylation sites:

amino acids 65-69,95-99,134-138,159-163,187-191,230-234,333-337

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 397-401

N-myristoylation sites:

amino acids 3-9,63-69,235-241,273-279,292-298,324-330

Leucine zipper pattern:

amino acids 371-393

FIGURE 259

CAGGCGGGCCCCCGCGCGGCAGGGCCCTGGACCCGCGCGGCTCCCGGGG**ATG**GTGAGCAAGGCGCTGCTGCGCCT

CGTGTCTGCCGTCAACCGCAGGAGGATGAAGCTGCTGCTGGGCATCGCCTTGCTGGCCTACGTCGCCTCTGTTTG

GGGCAACTTCGTTAATATGAGGTCTATCCAGGAAAATGGTGAACTAAAAATTGAAAGCAAGATTGAAGAGATGGT

TGAACCACTAAGAGAGAAAATCAGAGATTTAGAAAAAAGCTTTACCCAGAAATACCCACCAGTAAAGTTTTTATC

AGAAAAGGATCGGAAAAGAATTTTGATAACAGGAGGCGCAGGGTTCGTGGGCTCCCATCTAACTGACAAACTCAT

GATGGACGGCCACGAGGTGACCGTGGTGGACAATTTCTTCACGGGCAGGAAGAGAAACGTGGAGCACTGGATCGG

ACATGAGAACTTCGAGTTGATTAACCACGACGTGGTGGAGCCCCTCTACATCGAGGTTGACCAGATATACCATCT

GGCATCTCCAGCCTCCCCTCCAAACTACATGTATAATCCTATCAAGACATTAAAGACCAATACGATTGGGACATT

AAACATGTTGGGGCTGGCAAAACGAGTCGGTGCCCGTCTGCTCCTGGCCTCCACATCGGAGGTGTATGGAGATCC

TGAAGTCCACCCTCAAAGTGAGGATTACTGGGGCCACGTGAATCCAATAGGACCTCGGGCCTGCTACGATGAAGG

CAAACGTGTTGCAGAGACCATGTGCTATGCCTACATGAAGCAGGAAGGCGTGGAAGTGCGAGTGGCCAGAATCTT

CAACACCTTTGGGCCACGCATGCACATGAACGATGGGCGAGTAGTCAGCAACTTCATCCTGCAGGCGCTCCAGGG

GGAGCCACTCACGGTATACGGATCCGGGTCTCAGACAAGGGCGTTCCAGTACGTCAGCGATCTAGTGAATGGCCT

CGTGGCTCTCATGAACAGCAACGTCAGCAGCCCGGTCAACCTGGGGAACCCAGAAGAACACACAATCCTAGAATT

TGCTCAGTTAATTAAAAACCTTGTTGGTAGCGGAAGTGAAATTCAGTTTCTCTCCGAAGCCCAGGATGACCCACA

GAAAAGAAAACCAGACATCAAAAAAGCAAAGCTGATGCTGGGGTGGGAGCCCGTGGTCCCGCTGGAGGAAGGTTT

AAACAAAGCAATTCACTACTTCCGTAAAGAACTCGAGTACCAGGCAAATAATCAGTACATCCCCAAACCAAAGCC

TGCCAGAATAAAGAAAGGACGGACTCGCCACAGC**TGA**ACTCCTCACTTTTAGGACACAAGACTACCATTGTACAC

TTGATGGGATGTATTTTTGGCTTTTTTTTGTTGTCGTTTAAAGAAAGACTTTAACAGGTGTCATGAAGAACAAAC

TGGAATTTCATTCTGAAGCTTGCTTTAATGAAATGGATGTGCCTAAAAGCTCCCCTCAAAAAACTGCAGATTTTG

CCTTGCACTTTTTGAATCTCTCTTTTTATGTAAAATAGCGTAGATGCATCTCTGCGTATTTTCAAGTTTTTTTAT

CTTGCTGTGAGAGCATATGTTGTGACTGTCGTTGACAGTTTTATTTACTGGTTTCTTTGTGAAGCTGAAAAGGAA

CATTAAGCGGGACAAAAAATGCCGATTTTATTTATAAAAGTGGGTACTTAATAAATGAGTCGTTATACTATGCAT

AAAGAAAAATCCTAGCAGTATTGTCAGGTGGTGGTGCGCCGGCATTGATTTTAGGGCAGATAAAAGAATTCTGTG

TGAGAGCTTTATGTTTCTCTTTTAATTCAGAGTTTTTCCAAGGTCTACTTTTGAGTTGCAAACTTGACTTTGAAA

TATTCCTGTTGGTCATGATCAAGGATATTTGAAATCACTACTGTGTTTTGCTGCGTATCTGGGGCGGGGGCAGGT

TGGGGGGCACAAAGTTAACATATTCTTGGTTAACCATGGTTAAATATGCTATTTTAATAAAATATTGAAACTCA

FIGURE 260

MVSKALLRLVSAVNRRRMKLLLGIALLAYVASVWGNFVNMRSIQENGELKIESKIEEMVEPLR
EKIRDLEKSFTQKYPPVKFLSEKDRKRILITGGAGFVGSHLTDKLMMDGHEVTVVDNFFTGRK
RNVEHWIGHENFELINHDVVEPLYIEVDQIYHLASPASPPNYMYNPIKTLKTNTIGTLNMLGL
AKRVGARLLLASTSEVYGDPEVHPQSEDYWGHVNPIGPRACYDEGKRVAETMCYAYMKQEGVE
VRVARIFNTFGPRMHMNDGRVVSNFILQALQGEPLTVYGSGSQTRAFQYVSDLVNGLVALMNS
NVSSPVNLGNPEEHTILEFAQLIKNLVGSGSEIQFLSEAQDDPQKRKPDIKKAKLMLGWEPVV
PLEEGLNKAIHYFRKELEYQANNQYIPKPKPARIKKGRTRHS

Important features:

Signal peptide:

amino acids 1-32

N-glycosylation site:

amino acids 316-320

Tyrosine kinase phosphorylation site:

amino acids 235-244

N-myristoylation sites:

amino acids 35-41,101-107,383-389

Amidation sites:

amino acids 123-127,233-237

FIGURE 261

GCGTGGTGCGGGGGCGTGGGGAAATCGGGTTGCCCCAGCCGTTACTGGTCCGCGCAGTCAGGG
CATCCTCCGCATCCTCCACATCCTTCC**ATG**GCTCTGAAGAATAAATTCAGTTGTTTATGGATC
TTGGGTCTGTGTTTGGTAGCCACTACATCTTCCAAAATCCCATCCATCACTGACCCACACTTT
ATAGACAACTGCATAGAAGCCCACAACGAATGGCGTGGCAAAGTCAACCCTCCCGCGGCCGAC
ATGAAATACATGATTTGGGATAAAGGTTTAGCAAAGATGGCTAAAGCATGGGCAAACCAGTGC
AAATTTGAACATAATGACTGTTTGGATAAATCATATAAATGCTATGCAGCTTTTGAATATGTT
GGAGAAAATATCTGGTTAGGTGGAATAAAGTCATTCACACCAAGACATGCCATTACGGCTTGG
TATAATGAAACCCAATTTTATGATTTTGATAGTCTATCATGCTCCAGAGTCTGTGGCCATTAT
ACACAGTTAGTTTGGGCCAATTCATTTTATGTCGGTTGTGCAGTTGCAATGTGTCCTAACCTT
GGGGGAGCTTCAACTGCAATATTTGTATGCAACTACGGACCTGCAGGAAATTTTGCAAATATG
CCTCCTTACGCAAGAGGAGAATCTTGCTCTCTCTGCTCAAAAGAAGAGAAATGTGTAAAGAAC
CTCTGCAGGACTCCACAACTTATTATACCTAACCAAAATCCATTTCTGAAGCCAACGGGGAGA
GCACCTCAGCAGACAGCCTTTAATCCATTCAGCTTAGGTTTTCTTCTTCTGAGAATCTTT**TAA**
TGTCATTTATATACAAAGAAATTCTCAAATGTTAAAATAAAGGAATAGTTTATTGCTTAATA

FIGURE 262

MALKNKFSCLWILGLCLVATTSSKIPSITDPHFIDNCIEAHNEWRGKVNPPAADMKYMIWDKG

LAKMAKAWANQCKFEHNDCLDKSYKCYAAFEYVGENIWLGGIKSFTPRHAITAWYNETQFYDF

DSLSCSRVCGHYTQLVWANSFYVGCAVAMCPNLGGASTAIFVCNYGPAGNFANMPPYARGESC

SLCSKEEKCVKNLCRTPQLIIPNQNPFLKPTGRAPQQTAFNPFSLGFLLLRIF

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site:

amino acids 119-123

N-myristoylation sites:

amino acids 103-109,150-156,160-166,161-167,175-181

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1:

amino acids 136-156

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2:

amino acids 166-178

FIGURE 263

CGCCCTCCGACCCGCCCCGCGGCGCATTGTGGGATCTGTCGGCTTGTCAGGTGGTGGAGGAAA

AGGCGCTCCGTC**ATG**GGGATCCAGACGAGCCCCGTCCTGCTGGCCTCCCTGGGGGTGGGGCTG

GTCACTCTGCTCGGCCTGGCTGTGGGCTCCTACTTGGTTCGGAGGTCCCGCCGGCCTCAGGTC

ACTCTCCTGGACCCCAATGAAAGTACCTGCTACGACTGCTAGACAAGACGACTGTGAGCCAC

AACACCAAGAGGTTCCGCTTTGCCCTGCCCACCGCCCACCACACTCTGGGGCTGCCTGTGGGC

AAACATATCTACCTCTCCACCCGAATTGATGGCAGCCTGGTCATCAGGCCATACACTCCTGTC

ACCAGTGATGAGGATCAAGGCTATGTGGATCTTGTCATCAAGGTCTACCTGAAGGGTGTGCAC

CCCAAATTTCCTGAGGGAGGGAAGATGTCTCAGTACCTGGATAGCCTGAAGGTTGGGGATGTG

GTGGAGTTTCGGGGGCCAAGCGGGTTGCTCACTTACACTGGAAAAGGGCATTTTAACATTCAG

CCCAACAAGAAATCTCCACCAGAACCCCGAGTGGCGAAGAAACTGGGAATGATTGCCGGCGGG

ACAGGAATCACCCCAATGCTACAGCTGATCCGGGCCATCCTGAAAGTCCCTGAAGATCCAACC

CAGTGCTTTCTGCTTTTTGCCAACCAGACAGAAAAGGATATCATCTTGCGGGAGGACTTAGAG

GAACTGCAGGCCCGCTATCCCAATCGCTTTAAGCTCTGGTTCACTCTGGATCATCCCCCAAAA

GATTGGGCCTACAGCAAGGGCTTTGTGACTGCCGACATGATCCGGGAACACCTGCCCGCTCCA

GGGGATGATGTGCTGGTACTGCTTTGTGGGCCACCCCCAATGGTGCAGCTGGCCTGCCATCCC

AACTTGGACAAACTGGGCTACTCACAAAGATGCGATTCACCTAC**TGA**GCATCCTCCAGCTTC

CCTGGTGCTGTTCGCTGCAGTTGTTCCCCATCAGTACTCAAGCACTATAAGCCTTAGATTCCT

TTCCTCAGAGTTTCAGGTTTTTTTCAGTTACATCTAGAGCTGAAATCTGGATAGTACCTGCAGG

AACAATATTCCTGTAGCCATGGAAGAGGGCAAGGCTCAGTCACTCCTTGGATGGCCTCCTAAA

TCTCCCCGTGGCAACAGGTCCAGGAGAGGCCCATGGAGCAGTCTCTTCCATGGAGTAAGAAGG

AAGGGAGCATGTACGCTTGGTCCAAGATTGGCTAGTTCCTTGATAGCATCTTACTCTCACCTT

CTTTGTGTCTGTGATGAAAGGAACAGTCTGTGCAATGGGTTTTACTTAAACTTCACTGTTCAA

CCTATGAGCAAATCTGTATGTGTGAGTATAAGTTGAGCATAGCATACTTCCAGAGGTGGTNTT

ATGGAGATGGCAAGAAAGGAGGAAATGATTTCTTCAGATNTCAAAGGAGTCTGAAATATCATA

TTTCTGTGTGTGTCTCTCTCAGCCCCTGCCCAGGCTAGAGGGAAACAGCTACTGATAATCGAA

AACTGCTGTTTGTGGCANGAACCCCTGGCTGTGCAAATAAATGGGGCTGAGGCCCCTGTGTGA

TATTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURE 264

MGIQTSPVLLASLGVGLVTLLGLAVGSYLVRRSRRPQVTLLDPNEKYLLRLLDKTTVSHNTKR
FRFALPTAHHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSDEDQGYVDLVIKVYLKGVHPKFP
EGGKMSQYLDSLKVGDVVEFRGPSGLLTYTGKGHFNIQPNKKSPPEPRVAKKLGMIAGGTGIT
PMLQLIRAILKVPEDPTQCFLLFANQTEKDIILREDLEELQARYPNRFKLWFTLDHPPKDWAY
SKGFVTADMIREHLPAPGDDVLVLLCGPPPMVQLACHPNLDKLGYSQKMRFTY

Important features:

Signal peptide:

amino acids 1-26

N-glycosylation site:

amino acids 214-218

N-myristoylation sites:

amino acids 22-28,76-82,128-134,180-186

FIGURE 265

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAGA
ACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTATAG
AATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACTGCACCTCGGT
TCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCAC**ATG**TTGGGGGCCCGCCTCAGG
CTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCAATGCC
TCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCCAGGAAC
AGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGACCATCTAC
AGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGATGAGCAGA
AGATACCTCTGCATGGATTTCAGAGGCAACATTTTTGGATCACACTATTTCGACCCGGAGAAC
TGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCCTCAGTATCAC
TTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACCCACCCCCGTAC
TCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACCCCATACCACGG
CGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGTGCTGAAGCCCCGG
GCCCGGATGACCCCGGCCCCGGCCTCCTGTTCACAGGAGCTCCCGAGCGCCGAGGACAACAGC
CCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAACACGCACGCTGGGGGA
ACGGGCCCGGAAGGCTGCCGCCCCTTCGCCAAGTTCATC**TAG**GGTCGCTGG

FIGURE 266

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSSWGGLIHLYTATARNSYHLQIHKNGHVDG
APHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGYDV
YHSPQYHFLVSLGRAKRAFLPGMNPPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSERDP
LNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTHAGGTGPEGCRPFAKFI

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 267

GGCTGAGGGGAGGCCCGGAGCCTTTCTGGGGCCTGGGGGATCCTCTTGCACTGGTGGGTGGAGAGAAGCGCCTGC

AGCCAACCAGGGTCAGGCTGTGCTCACAGTTTCCTCTGGCGGCATGTAAAGGCTCCACAAAGGAGTTGGGAGTTC

AAATGAGGCTGCTGCGGACGGCCTGAGGATGGACCCCAAGCCCTGGACCTGCCGAGCGTGGCACTGAGGCAGCGG

CTGACGCTACTGTGAGGGAAAGAAGGTTGTGAGCAGCCCCGCAGGACCCCTGGCCAGCCCTGGCCCCAGCCTCTG

CCGGAGCCCTCTGTGGAGGCAGAGCCAGTGGAGCCCAGTGAGGCAGGGCTGCTTGGCAGCCACCGGCCTGCAACT

CAGGAACCCCTCCAGAGGCCATGGACAGGCTGCCCCGCTGACGGCCAGGGTGAAGCATGTGAGGAGCCGCCCCGG

AGCCAAGCAGGAGGGAAGAGGCTTTCATAGATTCTATTCACAAAGAATAACCACCATTTTGCAAGGACC**ATG**AGG

CCACTGTGCGTGACATGCTGGTGGCTCGGACTGCTGGCTGCCATGGGAGCTGTTGCAGGCCAGGAGGACGGTTTT

GAGGGCACTGAGGAGGGCTCGCCAAGAGAGTTCATTTACCTAAACAGGTACAAGCGGGCGGGCGAGTCCCAGGAC

AAGTGCACCTACACCTTCATTGTGCCCCAGCAGCGGGTCACGGGTGCCATCTGCGTCAACTCCAAGGAGCCTGAG

GTGCTTCTGGAGAACCGAGTGCATAAGCAGGAGCTAGAGCTGCTCAACAATGAGCTGCTCAAGCAGAAGCGGCAG

ATCGAGACGCTGCAGCAGCTGGTGGAGGTGGACGGCGGCATTGTGAGCGAGGTGAAGCTGCTGCGCAAGGAGAGC

CGCAACATGAACTCGCGGGTCACGCAGCTCTACATGCAGCTCCTGCACGAGATCATCCGCAAGCGGGACAACGCG

TTGGAGCTCTCCCAGCTGGAGAACAGGATCCTGAACCAGACAGCCGACATGCTGCAGCTGGCCAGCAAGTACAAG

GACCTGGAGCACAAGTACCAGCACCTGGCCACACTGGCCCACAACCAATCAGAGATCATCGCGCAGCTTGAGGAG

CACTGCCAGAGGGTGCCCTCGGCCAGGCCCGTCCCCCAGCCACCCCCCGCTGCCCCGCCCCGGGTCTACCAACCA

CCCACCTACAACCGCATCATCAACCAGATCTCTACCAACGAGATCCAGAGTGACCAGAACCTGAAGGTGCTGCCA

CCCCCTCTGCCCACTATGCCCACTCTCACCAGCCTCCCATCTTCCACCGACAAGCCGTCGGGCCCATGGAGAGAC

TGCCTGCAGGCCCTGGAGGATGGCCACGACACCAGCTCCATCTACCTGGTGAAGCCGGAGAACACCAACCGCCTC

ATGCAGGTGTGGTGCGACCAGAGACACGACCCCGGGGGCTGGACCGTCATCCAGAGACGCCTGGATGGCTCTGTT

AACTTCTTCAGGAACTGGGAGACGTACAAGCAAGGGTTTGGGAACATTGACGGCGAATACTGGCTGGGCCTGGAG

AACATTTACTGGCTGACGAACCAAGGCAACTACAAACTCCTGGTGACCATGGAGGACTGGTCCGGCCGCAAAGTC

TTTGCAGAATACGCCAGTTTCCGCCTGGAACCTGAGAGCGAGTATTATAAGCTGCGGCTGGGGCGCTACCATGGC

AATGCGGGTGACTCCTTTACATGGCACAACGGCAAGCAGTTCACCACCCTGGACAGAGATCATGATGTCTACACA

GGAAACTGTGCCCACTACCAGAAGGGAGGCTGGTGGTATAACGCCTGTGCCCACTCCAACCTCAACGGGGTCTGG

TACCGCGGGGGCCATTACCGGAGCCGCTACCAGGACGGAGTCTACTGGGCTGAGTTCCGAGGAGGCTCTTACTCA

CTCAAGAAAGTGGTGATGATGATCCGACCGAACCCCAACACCTTCCAC**TAA**GCCAGCTCCCCCTCCTGACCTCTC

GTGGCCATTGCCAGGAGCCCACCCTGGTCACGCTGGCCACAGCACAAAGAACAACTCCTCACCAGTTCATCCTGA

GGCTGGGAGGACCGGGATGCTGGATTCTGTTTTCCGAAGTCACTGCAGCGGATGATGGAACTGAATCGATACGGT

GTTTTCTGTCCCTCCTACTTTCCTTCACACCAGACAGCCCCTCATGTCTCCAGGACAGGACAGGACTACAGACAA

CTCTTTCTTTAAATAAATTAAGTCTCTACAATAAAAAAAA

FIGURE 268

MRPLCVTCWWLGLLAAMGAVAGQEDGFEGTEEGSPREFIYLNRYKRAGESQDKCTYTFIVPQQ
RVTGAICVNSKEPEVLLENRVHKQELELLNNELLKQKRQIETLQQLVEVDGGIVSEVKLLRKE
SRNMNSRVTQLYMQLLHEIIRKRDNALELSQLENRILNQTADMLQLASKYKDLEHKYQHLATL
AHNQSEIIAQLEEHCQRVPSARPVPQPPPAAPPRVYQPPTYNRIINQISTNEIQSDQNLKVLP
PPLPTMPTLTSLPSSTDKPSGPWRDCLQALEDGHDTSSIYLVKPENTNRLMQVWCDQRHDPGG
WTVIQRRLDGSVNFFRNWETYKQGFGNIDGEYWLGLENIYWLTNQGNYKLLVTMEDWSGRKVF
AEYASFRLEPESEYYKLRLGRYHGNAGDSFTWHNGKQFTTLDRDHDVYTGNCAHYQKGGWWYN
ACAHSNLNGVWYRGGHYRSRYQDGVYWAEFRGGSYSLKKVVMMIRPNPNTFH

Important features:

Signal peptide:
amino acids 1-22

N-glycosylation sites:
amino acids 164-168,192-196

cAMP- and cGMP-dependent protein kinase phosphorylation site:
amino acids 124-128

Tyrosine kinase phosphorylation sites:
amino acids 177-184,385-393,385-394,461-468

N-myristoylation sites:
amino acids 12-18,18-24,22-28,29-35,114-120,341-347,465-471,
473-479

Amidation site:
amino acids 373-377

Fibrinogen beta and gamma chains C-terminal domain signature:
amino acids 438-451

Fibrinogen beta and gamma chains C-terminal domain proteins:
amino acids 305-343, 365-402, 411-424, 428-458

Trehalase proteins:
amino acids 275-292

FIGURE 269

GCCGAGCTGAGCGGATCCTCAC**ATG**ACTGTGATCCGATTCTTTCCAGCGGCTTCTGCAACCAA
GCGGGTCTTACCCCCGGTCCTCCGCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCCCGA
GAGTCCCCGAATCCCCGCTCCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGACGGCCGGGGC
AGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTCAGGGCGGACCCGTGCAGTC
CAAGTCGCCGCGCTTTGCGTCCTGGGACGAGATGAATGTCCTGGCGCACGGACTCCTGCAGCT
CGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGCAGTCAGCTGAGCGCGCTGGAGCGGCG
CCTGAGCGCGTGCGGGTCCGCCTGTCAGGGAACCGAGGGGTCCACCGACCTCCCGTTAGCCCC
TGAGAGCCGGGTGGACCCTGAGGTCCTTCACAGCCTGCAGACACAACTCAAGGCTCAGAACAG
CAGGATCCAGCAACTCTTCCACAAGGTGGCCCAGCAGCAGCGGCACCTGGAGAAGCAGCACCT
GCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGACCACAAGCACCTAGACCATGAGGT
GGCCAAGCCTGCCCGAAGAAAGAGGCTGCCCGAGATGGCCCAGCCAGTTGACCCGGCTCACAA
TGTCAGCCGCCTGCACCGGCTGCCCAGGGATTGCCAGGAGCTGTTCCAGGTTGGGGAGAGGCA
GAGTGGACTATTTGAAATCCAGCCTCAGGGGTCTCCGCCATTTTTGGTGAACTGCAAGATGAC
CTCAGATGGAGGCTGGACAGTAATTCAGAGGCGCCACGATGGCTCAGTGGACTTCAACCGGCC
CTGGGAAGCCTACAAGGCGGGGTTTGGGGATCCCCACGGCGAGTTCTGGCTGGGTCTGGAGAA
GGTGCATAGCATCACGGGGGACCGCAACAGCCGCCTGGCCGTGCAGCTGCGGGACTGGGATGG
CAACGCCGAGTTGCTGCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCA
GCTCACTGCACCCGTGGCCGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGT
ACCCTTCTCCACTTGGGACCAGGATCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAGCCT
CTCTGGAGGCTGGTGGTTTGGCACCTGCAGCCATTCCAACCTCAACGGCCAGTACTTCCGCTC
CATCCCACAGCAGCGGCAGAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGCCGCTA
CTACCCGCTGCAGGCCACCACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCC**TAG**CG
TCCTGGCTGGGCCTGGTCCCAGGCCCACGAAAGACGGTGACTCTTGGCTCTGCCCGAGGATGT
GGCCGTTCCCTGCCTGGGCAGGGGCTCCAAGGAGGGGCCATCTGGAAACTTGTGGACAGAGAA
GAAGACCACGACTGGAGAAGCCCCCTTTCTGAGTGCAGGGGGGCTGCATGCGTTGCCTCCTGA
GATCGAGGCTGCAGGATATGCTCAGACTCTAGAGGCGTGGACCAAGGGGCATGGAGCTTCACT
CCTTGCTGGCCAGGGAGTTGGGGACTCAGAGGGACCACTTGGGCCAGCCAGACTGGCCTCAA
TGGCGGACTCAGTCACATTGACTGACGGGGACCAGGGCTTGTGTGGGTCGAGAGCGCCCTCAT
GGTGCTGGTGCTGTTGTGTGTAGGTCCCCTGGGGACACAAGCAGGCGCCAATGGTATCTGGGC
GGAGCTCACAGAGTTCTTGGAATAAAAGCAACCTCAGAACAC

FIGURE 270

MTVIRFFPAASATKRVLPPVLRVSSPRTWNPNVPESPRIPAPRLPKRMSGAPTAGAALMLCAA
TAVLLSAQGGPVQSKSPRFASWDEMNVLAHGLLQLGQGLREHAERTRSQLSALERRLSACGSA
CQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLFHKVAQQQRHLEKQHLRIQHLQS
QFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSRLHRLPRDCQELFQVGERQSGLFEIQ
PQGSPPFLVNCKMTSDGGWTVIQRRHDGSVDFNRPWEAYKAGFGDPHGEFWLGLEKVHSITGD
RNSRLAVQLRDWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPPSGLSVPFSTWDQ
DHDLRRDKNCAKSLSGGWWFGTCSHSNLNGQYFRSIPQQRQKLKKGIFWKTWRGRYYPLQATT
MLIQPMAAEAAS

Important features:
Signal peptide:
Amino acids 1-13

Transmembrane domain:
Amino acids 53-70

N-glycosylation site:
Amino acids 224-228

cAMP- and cGMP-dependent protein kinase phosphorylation sites:
Amino acids 46-50;118-122

N-myristoylation sites:
Amino acids 50-56;129-135;341-347;357-363

Fibrinogen beta and gamma chains C-terminal domain signature:
Amino acids 396-409

FIGURE 271

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGGG

AACAAG**ATG**GCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGCTG

CTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCGCTTCGGCTGAAGCATTTGACTCG

GTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTAC

CCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTT

GTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA

TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCGCT

GAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAACT

CTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACCTCTTCATGG

ACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCCAGAAATCCAG

TACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAAGCAAAATGTCC

TATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGAGAAAGTGATGGC

TTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCTTGTCCTCTCGGTG

ATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGCAGTATGTTCCCTCT

GAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAGCTAAACAGATATCCA

GCTTCTTCTCTTGTGGTTGTTAGATCTAAAACTGAAGATCATGAAGAAGCAGGGCCTCTACCT

ACAAAAGTGAATCTTGCTCATTCTGAAATT**TAA**GCATTTTTCTTTTAAAAGACAAGTGTAATA

GACATCTAAAATTCCACTCCTCATAGAGCTTTTAAAATGGTTTCATTGGATATAGGCCTTAAG

AAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 272

MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYPK
EEELYACQRGCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFAEL
RQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEIQYA
PHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVLSVMV
LLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAGPLPTK
VNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site:

amino acids 90-94

N-myristoylation sites:

amino acids 28-34,29-35,31-37,86-92

FIGURE 273

CCCACGCGTCCGAACCTCTCCAGCG**ATG**GGAGCCGCCCGCCTGCTGCCCAACCTCACTCTGTG
CTTACAGCTGCTGATTCTCTGCTGTCAAACTCAGTACGTGAGGGACCAGGGCGCCATGACCGA
CCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCACGT
GCAGGTCACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGCTCAT
AGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAGTACAT
CTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTGCGTGTT
CACGGAGATCGTGCTGGAGAACAACTATACGGCCTTCCAGAACGCCCGGCACGAGGGCTGGTT
CATGGCCTTCACGCGGCAGGGGCGGCCCGCCAGGCTTCCCGCAGCCGCCAGAACCAGCGCGA
GGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGCCGAGAAGCA
GAAGCAGTTCGAGTTTGTGGGCTCCGCCCCCACCCGCCGGACCAAGCGCACACGGCGGCCCCA
GCCCCTCACG**TAG**TCTGGGAGGCAGGGGGCAGCAGCCCCTGGGCCGCCTCCCCACCCCTTTCC
CTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCCGAGGGAGGACCC
TGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTGCCCCAGGGGCGGC
TGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACTGAGTGTCACCCTGA
TCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCCCGCTGAAAGGTCAGC
GACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAATCTGCTTCTCGGATCT
CCCTCAGTCTGCCCCCAGCCCCCAAACTCCTCCTGGCTAGACTGTAGGAAGGGACTTTTGTTT
GTTTGTTTGTTTCAGGAAAAAGAAAGGGAGAGAGAGGAAAATAGAGGGTTGTCCACTCCTCA
CATTCCACGACCCAGGCCTGCACCCCACCCCCAACTCCCAGCCCCGGAATAAAACCATTTTCC
TGC

FIGURE 274

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRIS
ATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLIGKPSGKSKDCVFTEIVLENN
YTAFQNARHEGWFMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEFVGS
APTRRTKRTRRPQPLT

Important features:

Signal peptide:

Amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 275

TATTTACCATATCAGATTCACATTCAGTCCTCAGCAAA**ATG**AAGGGCTCCATTTTCACTCTGT
TTTTATTCTCTGTCCTATTTGCCATCTCAGAAGTGCGGAGCAAGGAGTCTGTGAGACTCTGTG
GGCTAGAATACATACGGACAGTCATCTATATCTGTGCTAGCTCCAGGTGGAGAAGGCATCTGG
AGGGGATCCCTCAAGCTCAGCAAGCTGAGACAGGAAACTCCTTCCAGCTCCCACATAAACGTG
AGTTTTCTGAGGAAAATCCAGCGCAAAACCTTCCGAAGGTGGATGCCTCAGGGGAAGACCGTC
TTTGGGGTGGACAGATGCCCACTGAAGAGCTTTGGAAGTCAAAGAAGCATTCAGTGATGTCAA
GACAAGATTTACAAACTTTGTGTTGCACTGATGGCTGTTCCATGACTGATTTGAGTGCTCTTT
GC**TAA**GACAAGAGCAAATACCCAATGGGTGGCAGAGCTTTATCACATGTTTAATTACAGTGTT
TTACTGCCTGGTAGAACACTAATATTGTGTTATTAAAATGATGGCTTTTGGGTAGGCAAAACT
TCTTTTCTAAAAGGTATAGCTGAGCGGTTGAAACCACAGTGATCTCTATTTTCTCCCTTTGCC
AAGGTTAATGAACTGTTCTTTTCAAATTCTACTAATGCTTTGAAATTTCAAATGCTGCGCAAA
ATTGCAATAAAAATGCTATAAA

FIGURE 276

MKGSIFTLFLFSVLFAISEVRSKESVRLCGLEYIRTVIYICASSRWRRHLEGIPQAQQAETGN
SFQLPHKREFSEENPAQNLPKVDASGEDRLWGGQMPTEELWKSKKHSVMSRQDLQTLCCTDGC
SMTDLSALC

Important features:

Signal sequence:

amino acids 1-18

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 107-111

N-myristoylation sites:

amino acids 3-9,52-58,96-102,125-131

Insulin family signature:

amino acids 121-136

Insulin family proteins:

amino acids 28-46

FIGURE 277

```
GCAGCTGGTTACTGCATTTCTCCATGTGGCAGACAGAGCAAAGCCACAACGCTTTCTCTGCTGGATTAAAGACGG
CCCACAGACCAGAACTTCCACTATACTACTTAAAATTACATAGGTGGCTTGTCAAATTCAATTGATTAGTATTGT
AAAAGGAAAAAGAAGTTCCTTCTTACAGCTTGGATTCAACGGTCCAAAACAAAAATGCAGCTGCCATTAAAGTCT
CAGATGAACAAACTTCTACACTGATTTTTAAAATCAAGAATAAGGGCAGCAAGTTTCTGGATTCACTGAATCAAC
AGACACAAAAAGCTGGCAATATAGCAACTATGAAGAGAAAGCTACTAATAAAATTAACCCAACGCATAGAAGAC
TTTTTTTTCTCTTCTAAAAACAACTAAGTAAAGACTTAAATTTAAACACATCATTTTACAACCTCATTTCAAA**AT
G**AAGACTTTTACCTGGACCCTAGGTGTGCTATTCTTCCTACTAGTGGACACTGGACATTGCAGAGGTGGACAATT
CAAAATTAAAAAAATAAACCAGAGAAGATACCCTCGTGCCACAGATGGTAAAGAGGAAGCAAAGAAATGTGCATA
CACATTCCTGGTACCTGAACAAAGAATAACAGGGCCAATCTGTGTCAACACCAAGGGGCAAGATGCAAGTACCAT
TAAAGACATGATCACCAGGATGGACCTTGAAAACCTGAAGGATGTGCTCTCCAGGCAGAAGCGGGAGATAGATGT
TCTGCAACTGGTGGTGGATGTAGATGGAAACATTGTGAATGAGGTAAAGCTGCTGAGAAAGGAAAGCCGTAACAT
GAACTCTCGTGTTACTCAACTCTATATGCAATTATTACATGAGATTATCCGTAAGAGGGATAATTCACTTGAACT
TTCCCAACTGGAAAACAAAATCCTCAATGTCACCACAGAAATGTTGAAGATGGCAACAAGATACAGGGAACTAGA
GGTGAAATACGCTTCCTTGACTGATCTTGTCAATAACCAATCTGTGATGATCACTTTGTTGGAAGAACAGTGCTT
GAGGATATTTTCCCGACAAGACACCCATGTGTCTCCCCCACTTGTCCAGGTGGTGCCACAACATATTCCTAACAG
CCAACAGTATACTCCTGGTCTGCTGGGGAGGTAACGAGATTCAGAGGGATCCAGGTTATCCCAGAGATTTAATGCC
ACCACCTGATCTGGCAACTTCTCCCACCAAAAGCCCTTTCAAGATACCACCGGTAACTTTCATCAATGAAGGACC
ATTCAAAGACTGTCAGCAAGCAAAAGAAGCTGGGCATTCGGTCAGTGGGATTTATATGATTAAACCTGAAAACAG
CAATGGACCAATGCAGTTATGGTGTGAAAACAGTTTGGACCCTGGGGGTTGGACTGTTATTCAGAAAAGAACAGA
CGGCTCTGTCAACTTCTTCAGAAATTGGGAAAATTATAAGAAAGGGTTTGGAAACATTGACGGAGAATACTGGCT
TGGACTGGAAAATATCTATATGCTTAGCAATCAAGATAATTACAAGTTATTGATTGAATTAGAAGACTGGAGTGA
TAAAAAAGTCTATGCAGAATACAGCAGCTTTCGTCTGGAACCTGAAAGTGAATTCTATAGACTGCGCCTGGGAAC
TTACCAGGGAAATGCAGGGGATTCTATGATGTGGCATAATGGTAAACAATTCACCACACTGGACAGAGATAAAGA
TATGTATGCAGGAAACTGCGCCCACTTTCATAAAGGAGGCTGGTGGTACAATGCCTGTGCACATTCTAACCTAAA
TGGAGTATGGTACAGAGGAGGCCATTACAGAAGCAAGCACCAAGATGGAATTTTCTGGGCCGAATACAGAGGCGG
GTCATACTCCTTAAGAGCAGTTCAGATGATGATCAAGCCTATTGAC**TGA**AGAGAGACACTCGCCAATTTAAATGA
CACAGAACTTTGTACTTTTCAGCTCTTAAAAATGTAAATGTTACATGTATATTACTTGGCACAATTTATTTCTAC
ACAGAAAGTTTTTAAAATGAATTTTACCGTAACTATAAAAGGGAACCTATAAATGTAGTTTCATCTGTCGTCAAT
TACTGCAGAAAATTATGTGTATCCACAACCTAGTTATTTTAAAAATTATGTTGACTAAATACAAAGTTTGTTTTC
TAAAATGTAAATATTTGCCACAATGTAAAGCAAATCTTAGCTATATTTTAAATCATAAATAACATGTTCAAGATA
CTTAACAATTTATTTAAAATCTAAGATTGCTCTAACGTCTAGTGAAAAAAATATTTTTTAAATTTCAGCCAAATA
ATGCATTTTATTTTATAAAAATACAGACAGAAAATTAGGGAGAAACTTCTAGTTTTGCCAATAGAAAATGTTCTT
CCATTGAATAAAAGTTATTTCAAATTGAATTTGTGCCTTTCACACGTAATGATTAAATCTGAATTCTTAATAATA
TATCCTATGCTGATTTTCCCAAAACATGACCCATAGTATTAAATACATATCATTTTTAAAAATAAAAAAAAACCC
AAAAATAATGCATGCATAATTTAAATGGTCAATTTATAAAGACAAATCTATGAATGAATTTTTCAGTGTTATCTT
CATATGATATGCTGAACACCAAAATCTCCAGAAATGCATTTTATGTAGTTCTAAAATCAGCAAAATATTGGTATT
ACAAAAATGCAGAATATTTAGTGTGCTACAGATCTGAATTATAGTTCTAATTTATTATTACTTTTTTTTCTAATTT
ACTGATCTTACTACTACAAAGAAAAAAAAACCCAACCCATCTGCAATTCAAATCAGAAAGTTTGGACAGCTTTAC
AAGTATTAGTGCATGCTCAGAACAGGTGGGACTAAAACAAACTCAAGGAACTGTTGGCTGTTTTCCCGATACTGA
GAATTCAACAGCTCCAGAGCAGAAGCCACAGGGGCATAGCTTAGTCCAAACTGCTAATTTCATTTTACAGTGTAT
GTAACGCTTAGTCTCACAGTGTCTTTAACTCATCTTTGCAATCAACAACTTTACTAGTGACTTTCTGGAACAATT
TCCTTTCAGGAATACATATTCACTGCTTAGAGGTGACCTTGCCTTAATATATTTGTGAAGTTAAAATTTTAAAGA
TAGCTCATGAAACTTTTGCTTAAGCAAAAGAAAACCTCGAATTGAAATGTGTGAGGCAAACTATGCATGGGAAT
AGCTTAATGTGAAGATAATCATTTGGACAACTCAAATCCATCAACATGACCAATGTTTTTCATCTGCCACATCTC
AAAATAAAACTTCTGGTGAAACAAATTAAACAAAATATCCAAACCTCAAAAAAAA
```

FIGURE 278

MKTFTWTLGVLFFLLVDTGHCRGGQFKIKKINQRRYPRATDGKEEAKKCAYTFLVPEQRITGP
ICVNTKGQDASTIKDMITRMDLENLKDVLSRQKREIDVLQLVVDVDGNIVNEVKLLRKESRNM
NSRVTQLYMQLLHEIIRKRDNSLELSQLENKILNVTTEMLKMATRYRELEVKYASLTDLVNNQ
SVMITLLEEQCLRIFSRQDTHVSPPLVQVVPQHIPNSQQYTPGLLGGNEIQRDPGYPRDLMPP
PDLATSPTKSPFKIPPVTFINEGPFKDCQQAKEAGHSVSGIYMIKPENSNGPMQLWCENSLDP
GGWTVIQKRTDGSVNFFRNWENYKKGFGNIDGEYWLGLENIYMLSNQDNYKLLIELEDWSDKK
VYAEYSSFRLEPESEFYRLRLGTYQGNAGDSMMWHNGKQFTTLDRDKDMYAGNCAHFHKGGWW
YNACAHSNLNGVWYRGGHYRSKHQDGIFWAEYRGGSYSLRAVQMMIKPID

Important features:

Signal sequence:

Amino acids 1-23

N-glycosylation sites:

Amino acids 160-164;188-192

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 120-124

Tyrosine kinase phosphorylation sites:

Amino acids 173-180;387-396

N-myristoylation sites:

Amino acids 70-76;110-116;232-238,343-349;400-406;467-473;
475-487

Fibrinogen beta and gamma chains C-terminal domain signature:

Amino acids 440-453

FIGURE 279

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGGC

CCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGCCG

CCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAAGCC

CGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAGCGGA

GGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCAGCAAAAGAAGCGGTGGTGGTGGGCGTCG

TGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCGAGCGCG

AGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGCGACAAAA

ACAAGTTAAATGTCTTTTCCCGGGTCAAACTCTTCGGCTCCAAGAAGAGGCGCAGAAGAAGAC

CAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACCACTTGCAGC

TGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACTCTGTTTAACC

TCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCTGTACTTGGCAA

TGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCAAATTCAAAGAAT

CAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAGCAGCAGTCAGGCC

GAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAACCATGTGAAGAAGA

ACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGTACAAGGAGCCATCAC

TGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACCAAGAGCAGAAGTGTCT

CTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAACG**TAG**CCAGTGAGGGCAA

AAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAATTCTTCTAGCAGTCCTTCA

CCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGGCAGAGTTCACTATTCTATC

TGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 280

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFGSKKRRRRRPE
PQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLAMN
SEGYLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVKKNK
PAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMSHNEST

Important Features:

N-glycosylation site:

Amino acids 242-246

Glycosaminoglycan attachment sites:

Amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site:

Amino acids 93-100

N-myristoylation sites:

Amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop):

Amino acids 231-239

HBGF/FGF family proteins:

Amino acids 78-94, 102-153

FIGURE 281

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGGA
CTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGGCA
ACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCTCCA
TTTAGAGTGTAGCAAAGGAAAAAACACCAAGGTTGGGTTCCTTCCTGACATTGGCAGTGCCCC
AGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTACAAGAG
TGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAAACCACG
TCTTGGAAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTGGAGAGGA
GGGAAAGGGGACGTTTTCAATAGGAGGCAAAACTCGAGGGTGGGATCCACTGAGGAGTACATA
GGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCTGTGGAGGGG
GGTACGTGAGGGGGGGGGTCTGGGGCTTATCCTCAGGTCCTGTGGGTGGGGCAGCGAGTCGGGG
CCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAGCGCGCTCCGGG
CGCCTGCCGGTTTGGGGGTGTCTCCTCCCGGGGCGCT**ATG**GCGGCGCTGGCCAGTAGCCTGAT
CCGGCAGAAGCGGGAGGTCCGCGAGCCCGGGGGCAGCCGGCCGGTGTCGGCGCAGCGGCGCGT
GTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTGCTGTCCAAGGTGCG
ACTGTGCGGGGGGCGGCCCGCGCGGCCGGACCGCGGCCCGGAGCCTCAGCTCAAAGGCATCGT
CACCAAACTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCGACGGAAGCATCCAGGG
CACCCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCTGTGGGCCTCCGTGTGGT
CACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGCTGAGGGACTGCTCTACAG
TTCGCCGCATTTCACAGCTGAGTGTCGCTTTAAGGAGTGTGTCTTTGAGAATTACTACGTCCT
GTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGCCTGGTACCTCGGCCTGGACAA
GGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAAGGCAGCTGCCCACTTTCTGCC
CAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCCACAGTGTCCCCGAGGCCTCCCC
TTCCAGTCCCCCTGCCCCC**TGA**AATGTAGTCCCTGGACTGGAGGTTCCCTGCACTCCCAGTGA
GCCAGCCACCACCACAACCTGT

FIGURE 282

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDRG
PEPQLKGIVTKLFCRQGFYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHYMA
MNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNRVKK
TKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

Important features:

Tyrosine kinase phosphorylation site:

Amino acids 199-207

N-myristoylation sites:

Amino acids 54-60; 89-95; 131-137

HBGF/FGF family signature:

Amino acids 131-155

FIGURE 283

ATGGCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTGG
GACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACGGC
AACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGGCGC
CAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTTGCAA
ATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCTTCAAC
CTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTATATAGCC
ATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTTTAAAGAA
TCTGTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGGAATCTGGT
AGAGCCTGGTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGAGTAAAGAAA
ACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCGAGAACCATCT
TTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAAGCACAAGTGCG
TCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

FIGURE 284

MAAAIASGLIRQKRQAREQHWDRPSASRRRSSPSKNRGLCNGNLVDIFSKVRIFGLKKRRLRR
QDPQLKGIVTRLYCRQGYYLQMHPDGALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLYIA
MNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNRVKK
TKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKSKTT

Important features:

N-glycosylation sites:

Amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site:

Amino acids 199-207

N-myristoylation sites:

Amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins:

Amino acids 104-155, 171-198

FIGURE 285

```
CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAACTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAAACTACCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCCAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTGCCCGCCGTGAGTGAGCTCTCACCCCAGTCAGCCAAATGAGCCTCTTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTCCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTCATGCCACAATTCACAGAAGCTGTGAGTCCTTCAGTGCTACCCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGGAAGACCTTATTCGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCT
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCT
TCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTAT
TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTC
GATACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGT
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCT
ATTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAATACATGGATATTTTTATGTACA
GAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAAATATTTTGCTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAAGA
ATGTGGCTATTTTGGGGAGAAAATTAAAAAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC
```

FIGURE 286

MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPRF
PHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWCGS
GTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLLNNA
ITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLYSCTP
RNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQLRPKT
GVRGLHKSLTDVALEHHEECDCVCRGSTGG

Important features:

signal sequence:

Amino acids 1-14

N-glycosylation sites:

Amino acids 25-29;55-59;254-258

N-myristoylation sites:

Amino acids 15-21;117-123;127-133;281-287;282-288;319-325

Amidation site:

Amino acids 229-233

FIGURE 287

CAGCGCTGACTGCGCCGCGGAGAAAGCCAGTGGGAACCCAGACCCATAGGAGACCCGCGTCCC

CGCTCGGCCTGGCCAGGCCCCGCGCT**ATG**GAGTTCCTCTGGGCCCCTCTCTTGGGTCTGTGCT

GCAGTCTGGCCGCTGCTGATCGCCACACCGTCTTCTGGAACAGTTCAAATCCCAAGTTCCGGA

ATGAGGACTACACCATACATGTGCAGCTGAATGACTACGTGGACATCATCTGTCCGCACTATG

AAGATCACTCTGTGGCAGACGCTGCCATGGAGCAGTACATACTGTACCTGGTGGAGCATGAGG

AGTACCAGCTGTGCCAGCCCCAGTCCAAGGACCAAGTCCGCTGGCAGTGCAACCGGCCCAGTG

CCAAGCATGGCCCGGAGAAGCTGTCTGAGAAGTTCCAGCGCTTCACACCTTTCACCCTGGGCA

AGGAGTTCAAAGAAGGACACAGCTACTACTACATCTCCAAACCCATCCACCAGCATGAAGACC

GCTGCTTGAGGTTGAAGGTGACTGTCAGTGGCAAAATCACTCACAGTCCTCAGGCCCATGACA

ATCCACAGGAGAAGAGACTTGCAGCAGATGACCCAGAGGTGCGGGTTCTACATAGCATCGGTC

ACAGTGCTGCCCCACGCCTCTTCCCACTTGCCTGGACTGTGCTGCTCCTTCCACTTCTGCTGC

TGCAAACCCCG**TGA**AGGTGTGTGCCACACCTGGCCTTAAAGAGGGACAGGCTGAAGAGAGGGA

CAGGCACTCCAAACCTGTCTTGGGGCCACTTTCAGAGCCCCCAGCCCTGGGAACCACTCCCAC

CACAGGCATAAGCTATCACCTAGCAGCCTCAAAACGGGTCAATATTAAGGTTTTCAACCGGAA

GGAGGCCAACCAGCCCGACAGTGCCATCCCCACCTTCACCTCGGAGGGATGGAGAAAGAAGTG

GAGACAGTCCTTTCCCACCATTCCTGCCTTTAAGCCAAAGAAACAAGCTGTGCAGGCATGGTC

CCTTAAGGCACAGTGGGAGCTGAGCTGGAAGGGGCCACGTGGATGGGCAAAGCTTGTCAAAGA

TGCCCCCTTCAGGAGAGAGCCAGGATGCCCAGATGAACTGACTGAAGGAAAAGCAAGAAACAG

TTTCTTGCTTGGAAGCCAGGTACAGGAGAGGCAGCATGCTTGGGCTGACCCAGCATCTCCCAG

CAAGACCTCATCTGTGGAGCTGCCACAGAGAAGTTTGTAGCCAGGTACTGCATTCTCTCCCAT

CCTGGGGCAGCACTCCCCAGAGCTGTGCCAGCAGGGGGGCTGTGCCAACCTGTTCTTAGAGTG

TAGCTGTAAGGGCAGTGCCCATGTGTACATTCTGCCTAGAGTGTAGCCTAAAGGGCAGGGCCC

ACGTGTATAGTATCTGTATATAAGTTGCTGTGTGTCTGTCCTGATTTCTACAACTGGAGTTTT

TTTATACAATGTTCTTTGTCTCAAAATAAAGCAATGTGTTTTTCGG

FIGURE 288

MEFLWAPLLGLCCSLAAADRHTVFWNSSNPKFRNEDYTIHVQLNDYVDIICPHYEDHSADAAM

EQYILYLVEHEEYQLCQPQSKDQVRWQCNRPSAKHGPEKLSEKFQRFTPFTLGKEFKEGHSYY

YISKPIHQHEDRCLRLKVTVSGKITHSPQAHDNPQEKRLAADDPEVRVLHSIGHSAAPRLFPL

AWTVLLLPLLLLQTP

Important features:

Signal sequence:

Amino acids 1-17

N-glycosylation site:

Amino acids 26-30

Tyrosine kinase phosphorylation site:

Amino acids 118-127

N-myristoylation site:

Amino acids 10-16

FIGURE 289

CGGACGCGTGGGCGGACGCGTGGGCGGCCCACGGCGCCCGCGGGCTGGGGCGGTCGCTTCTTC
CTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAG**ATG**GCCCCATGGCCCCCGAAGGGC
CTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGGCTC
CAGCCCTCTCCACCTCCCCAGTCTTCTCCCCGCCTCAGCCCCATCCGTGTCATACCTGCCGG
GGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAGGTGGA
AACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTGGTAGAG
GTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCTGAGTGAG
GAGCTGGTGGAGAGCTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCCAGTGGCTG
TGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGCCTTCCCTGT
CCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGGGACACGAGGG
GGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCCAGTGTGGCCTT
GGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTCGGCTTGTTTTGGCCCCTGT
GCCCGATGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGGCTGGGCCCTGCAT
CACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACTGTGGAGCTGACCAA
TTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAGGCCTGCCTAGGCTGC
ATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCAGCAGGTGGGCTCCAAG
TGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGAACAAGCAGTGTGAAAAC
ACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAGATGGAAGGCATCTGTGTG
AAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGACAGAAGACGAGTTGGTGGTG
CTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCACGCTGGCTGCTAAGGGCGAC
TTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATGACTGGCTACTGGTTGTCAGAG
CGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGA**TAA**TCGCGGCCACCACCTGTAGGA
CCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCCTCCTGCTGGACACTCAGGACAGC
TTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTGCCTTACAGAGCAGCCCAGGTACCC
AGGCCCGGGCAGACAAGGCCCCTGGGGTAAAAGTAGCCCTGAAGGTGGATACCATGAGCTCT
TCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGAATTTCAAAAGTTTTTCCTTAATGGTG
GCTGCTAGAGCTTTGGCCCCTGCTTAGGATTAGGTGGTCCTCACAGGGGTGGGGCCATCACAG
CTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGTTCTGTGTTCACCACATCCCCACACCCCA
TTGCCACTTATTTATTCATCTCAGGAAATAAAGAAGGTCTTGGAAAGTTAAAAAAAAAAAAAA
AAAAAAAAAA

FIGURE 290

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLERT
IRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQQE
APDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAGYGG
EACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDECGTE
GANCGADQFCVNTEGSYECRDCAKACLGCMGAPGRCKKCSPGYQQVGSKCLDVDECETEVCP
GENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFGIIICAL
ATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Important features:

Signal sequence:

Amino acids 1-29

Transmembrane domain:

Amino acids 342-392

N-glycosylation sites:

Amino acids 79-83;205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 290-294

Aspartic acid and asparagine hydroxylation site:

Amino acids 321-333

EGF-like domain cysteine pattern signature:

Amino acids 181-193

FIGURE 291

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGACCCAC
GCGTCCGAACACAGGTCCTTGTTGCTGCAGAGAAGCAGTTGTTTTGCTGGAAGGAGGGAGTGCGCGGGCTGCCCC
GGGCTCCTCCCTGCCGCCTCCTCTCAGTGGATGGTTCCAGGCACCCTGTCTGGGGCAGGGAGGGCACAGGCCTGC
ACATCGAAGGTGGGGTGGGACCAGGCTGCCCCTCGCCCCAGCATCCAAGTCCTCCCTTGGGCGCCCGTGGCCCTG
CAGACTCTCAGGGCTAAGGTCCTCTGTTGCTTTTTGGTTCCACCTTAGAAGAGGCTCCGCTTGACTAAGAGTAGC
TTGAAGGAGGCACC**ATG**CAGGAGCTGCATCTGCTCTGGTGGGCGCTTCTCCTGGGCCTGGCTCAGGCCTGCCCTG
AGCCCTGCGACTGTGGGGAAAAGTATGGCTTCCAGATCGCCGACTGTGCCTACCGCGACCTAGAATCCGTGCCGC
CTGGCTTCCCGGCCAATGTGACTACACTGAGCCTGTCAGCCAACCGGCTGCCAGGCTTGCCGGAGGGTGCCTTCA
GGGAGGTGCCCCTGCTGCAGTCGCTGTGGCTGGCACACAATGAGATCCGCACGGTGGCCGCCGGAGCCCTGGCCT
CTCTGAGCCATCTCAAGAGCCTGGACCTCAGCCACAATCTCATCTCTGACTTTGCCTGGAGCGACCTGCACAACC
TCAGTGCCCTCCAATTGCTCAAGATGGACAGCAACGAGCTGACCTTCATCCCCCGCGACGCCTTCCGCAGCCTCC
GTGCTCTGCGCTCGCTGCAACTCAACCACAACCGCTTGCACACATTGGCCGAGGGCACCTTCACCCCGCTCACCG
CGCTGTCCCACCTGCAGATCAACGAGAACCCCTTCGACTGCACCTGCGGCATCGTGTGGCTCAAGACATGGGCCC
TGACCACGGCCGTGTCCATCCCGGAGCAGGACAACATCGCCTGCACCTCACCCCATGTGCTCAAGGGTACACCGC
TGAGCCGCCTGCCGCCACTGCCATGCTCGGCGCCCTCAGTGCAGCTCAGCTACCAACCCAGCCAGGATGGTGCCG
AGCTGCGGCCTGGTTTTGTGCTGGCACTGCACTGTGATGTGGACGGGCAGCCGGCCCCTCAGCTTCACTGGCACA
TCCAGATACCCAGTGGCATTGTGGAGATCACCAGCCCCAACGTGGGCACTGATGGGCGTGCCCTGCCTGGCACCC
CTGTGGCCAGCTCCCAGCCGCGCTTCCAGGCCTTTGCCAATGGCAGCCTGCTTATCCCCGACTTTGGCAAGCTGG
AGGAAGGCACCTACAGCTGCCTGGCCACCAATGAGCTGGGCAGTGCTGAGAGCTCAGTGGACGTGGCACTGGCCA
CGCCCGGTGAGGGTGGTGAGGACACACTGGGGCGCAGGTTCCATGGCAAAGCGGTTGAGGGAAAGGGCTGCTATA
CGGTTGACAACGAGGTGCAGCCATCAGGGCCGGAGGACAATGTGGTCATCATCTACCTCAGCCGTGCTGGGAACC
CTGAGGCTGCAGTCGCAGAAGGGGTCCCTGGGCAGCTGCCCCCAGGCCTGCTCCTGCTGGGCCAAAGCCTCCTCC
TCTTCTTCTTCCTCACCTCCTTC**TAG**CCCCACCCAGGGCTTCCCTAACTCCTCCCCTTGCCCCTACCAATGCCCC
TTTAAGTGCTGCAGGGGTCTGGGGTTGGCAACTCCTGAGGCCTGCATGGGTGACTTCACATTTTCCTACCTCTCC
TTCTAATCTCTTCTAGAGCACCTGCTATCCCCAACTTCTAGACCTGCTCCAAACTAGTGACTAGGATAGAATTTG
ATCCCCTAACTCACTGTCTGCGGTGCTCATTGCTGCTAACAGCATTGCCTGTGCTCTCCTCTCAGGGGCAGCATG
CTAACGGGGCGACGTCCTAATCCAACTGGGAGAAGCCTCAGTGGTGGAATTCCAGGCACTGTGACTGTCAAGCTG
GCAAGGGCCAGGATTGGGGGAATGGAGCTGGGGCTTAGCTGGGAGGTGGTCTGAAGCAGACAGGGAATGGGAGAG
GAGGATGGGAAGTAGACAGTGGCTGGTATGGCTCTGAGGCTCCCTGGGGCCTGCTCAAGCTCCTCCTGCTCCTTG
CTGTTTTCTGATGATTTGGGGGCTTGGGAGTCCCTTTGTCCTCATCTGAGACTGAAATGTGGGGATCCAGGATGG
CCTTCCTTCCTCTTACCCTTCCTCCCTCAGCCTGCAACCTCTATCCTGGAACCTGTCCTCCCTTTCTCCCCAACT
ATGCATCTGTTGTCTGCTCCTCTGCAAAGGCCAGCCAGCTTGGGAGCAGCAGAGAAATAAACAGCATTTCTGATG
CCAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCT

FIGURE 292

MQELHLLWWALLLGLAQACPEPCDCGEKYGFQIADCAYRDLESVPPGFPANVTTLSLSANRLP

GLPEGAFREVPLLQSLWLAHNEIRTVAAGALASLSHLKSLDLSHNLISDFAWSDLHNLSALQL

LKMDSNELTFIPRDAFRSLRALRSLQLNHNRLHTLAEGTFTPLTALSHLQINENPFDCTCGIV

WLKTWALTTAVSIPEQDNIACTSPHVLKGTPLSRLPPLPCSAPSVQLSYQPSQDGAELRPGFV

LALHCDVDGQPAPQLHWHIQIPSGIVEITSPNVGTDGRALPGTPVASSQPRFQAFANGSLLIP

DFGKLEEGTYSCLATNELGSAESSVDVALATPGEGGEDTLGRRFHGKAVEGKGCYTVDNEVQP

SGPEDNVVIIYLSRAGNPEAAVAEGVPGQLPPGLLLLGQSLLLFFFLTSF

Important features:

Signal peptide:
amino acids 1-18

Transmembrane domain:
amino acids 403-418

N-glycosylation sites:
Amino acids 51-55,120-124,309-313

Tyrosine kinase phosphorylation site:
amino acids 319-326

N-myristoylation sites:
amino acids 14-20,64-70,92-98,218-224,294-300,323-329,334-340,
350-356,394-400

Amidation site:
amino acids 355-359

Leucine Rich Repeat:
amino acids 51-74,75-98, 99-122,123-146,147-170

Leucine rich repeat C-terminal domain:
amino acids 180-230

FIGURE 293

```
ACTTGGAGCAAGCGGCGGCGGCGGAGACAGAGGCAGAGGCAGAAGCTGGGGCTCCGTCCTCGCCTCCCACGAGCG
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCTGCC
TCGCTTCCCAGGCGCCGGCGGCTGCAGCCTTGCCCCTCTTGCTCGCCTTGAAAATGGAAAAGATGCTCGCAGGCT
GCTTTCTGCTGATCCTCGGACAGATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGCTCACGTGGGAGGTCCATCT
CTAGGGGCAGACACGCTCGGACCCACCCGCAGACGGCCCTTCTGGAGAGTTCCTGTGAGAACAAGCGGGCAGACC
TGGTTTTCATCATTGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTCATCGTGGACA
TCTTGCAATTCTTGGACATTGGTCCTGATGTCACCCGAGTGGGCCTGCTCCAATATGGCAGCACTGTCAAGAATG
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGCTGTCAAGAGGATGCGGCATCTGTCCACGG
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCCGGCCCCTGA
GGGAGAATGTGCCACGGGTCATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA
AGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG
GGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTTCAGCCAGATTGAGACGCTGACCTCCGTGTTCC
AGAAGAAGTTGTGCACGGCCCACATGTGCAGCACCCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG
GCTCATACGTCTGCAGGTGCAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT
GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAAAACCACGGATGTGAAC
ATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAAA
CGTGCACAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCGTCAACATGGAGGAGAGCT
ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCCTTCGTCTGCCAGTGCTCAGAAGGCTTCC
TCATCAACGAGGACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGCTGAGTGACCATGGTTGTGAATACTCCT
GTGTCAACATGGACAGATCCTTTGCCTGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAGACGTGTG
CAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAACATTCGTGTGTAAGCAGTGAAGATTCGTTTGTGT
GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGGCTCG
CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAAATCAACCCACCATGGCTGCGAACACATTTGTGTTA
ATAATGGGAATTCCTACATCTGCAAATGCTCAGAGGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT
GCACTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTTGAGGTCG
TGAAGCAGTTTGTCACTGGAATTATAGATTCCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT
ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCC
ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTTACCCAAG
GAGAAGGGGCCAGGCCCCCTTTCCACAAGGGTGCCCAGAGCAGCCATTGTGTTCACCGACGGACGGGCTCAGGATG
ACGTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG
AGGAGGAACTACAAGAGATTGCCTCTGAGCCCACAAACAAGCATCTCTTCTATGCCGAAGACTTCAGCACAATGG
ATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG
CAGGGGAACTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT
CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT
CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGT
TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC
TGGAAAATCGCCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT
GAACGCAGTGCAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAACAATCAGTACTGA
GAAACCTGGTTTGCCACAGAACAAAGACAAGAAGTATACACTAACTTGTATAAATTTATCTAGGAAAAAAATCCT
TCAGAATTCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAATATACTGTGGACAC
AACTTGCTTCTGCCTCATCCTGCCTTAGTGTGCAATCTCATTTGACTATCGATAAAGTTTGCACAGTCTTACTT
CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG
CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAAAATTCACCACTTCAG
```

FIGURE 294

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVFIID
SSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAVKR
MRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARDTGI
LIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHNCAHF
CINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYALAEDG
KRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGCEHECVN
MEESYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLINEDLKTC
SRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCEHSCVSSED
SFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLAEDGKRCRRK
DVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVIDGSKSLGEENF
EVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVAHMKYMGKGSMT
GLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKANGITMYAVGVGKA
IEEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDSPAGELPKTVQQPT
ESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPLEEKHDQCKCENLIM
FQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Important features:
Signal sequence:
Amino acids 1-23

N-glycosylation site:
Amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation sites:
Amino acids 115-119;606-610;892-896

N-myristoylation sites:
Amino acids 133-139;258-264;299-305;340-346;453-459;494-500;
639-645;690-694;
752-758;792-798

Amidation sites:
Amino acids 314-318;560-564;601-605

Aspartic acid and asparagine hydroxylation sites:
Amino acids 253-265;294-306;335-347;376-388;417-429;
458-470;540-552;581-593

FIGURE 295

GGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCGCAGCGCTACCCGC
C**ATG**CGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTGCTGCTGCCGCCCGC
GCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCTGGTGGACAAGTTTAA
CCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACACGGCTTGGGAGGAAAA
GACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATCCTGGAGGGGCTGTGCGA
GAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGAGCACCTGGAGGCCTGGTG
GCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTTGTGTGAAGACACTGAAAGT
GTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGCCAGGGCGGATCCCAGAGGCC
CTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACAGGGCGACGGGTCCTGCCGGTG
CCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGGACGGCTACTTCAGCTCGCTCCG
GAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCCTGCAAGACGTGCTCGGGCCTGAC
CAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCTGGACGAGGGCGCCTGTGTGGATGT
GGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTGCGCAGTTCTGTAAGAACGCCAACGG
CTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTGGGCTGCACAGGGGAAGGCCCAGGAAA
CTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCACGGACAGTGTGCAGATGTGGACGAGTG
CTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACGAAACTGCTACAATACTCCAGGGAGCTA
CGTCTGTGTGTGTCCTGACGGCTTCGAAGAAACGGAAGATGCCTGTGTGCCGCCGGCAGAGGC
TGAAGCCACAGAAGGAGAAAGCCCGACACAGCTGCCCTCCCGCGAAGACCTG**TAA**TGTGCCGG
ACTTACCCTTTAAATTATTCAGAAGGATGTCCCGTGGAAAATGTGGCCCTGAGGATGCCGTCT
CCTGCAGTGGACAGCGGCGGGGAGAGGCTGCCTGCTCTCTAACGGTTGATTCTCATTTGTCCC
TTAAACAGCTGCATTTCTTGGTTGTTCTTAAACAGACTTGTATATTTTGATACAGTTCTTTGT
AATAAAATTGACCATTGTAGGTAATCAGGAGGAAAAAAAA

FIGURE 296

MRLPRRAALGLLPLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEK
TLSKYESSEIRLLEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKV
CCSPGTYGPDCLACQGGSQRPCSGNGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLR
NETHSICTACDESCKTCSGLTNRDCGECEVGWVLDEGACVDVDECAAEPPPCSAAQFCKNANG
SYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKTCVRKNENCYNTPGSY
VCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL

Important features:

Signal peptide:
Amino acids 1-24

N-glycosylation sites:
Amino acids 190-194;251-255

Glycosaminoglycan attachment sites:
Amino acids 149-153;155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 26-30

Tyrosine kinase phosphorylation site:
Amino acids 303-310

N-myristoylation sites:
Amino acids 44-50;54-60;55-61;81-87;150-156;158-164;164-170;
252-258;313-319

Aspartic acid and asparagine hydroxylation site:
Amino acids 308-320

EGF-like domain cysteine pattern signature:
Amino acids 166-178

Leucine zipper pattern:
Amino acids 94-116

FIGURE 297

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAGAAAGAAGAGGAAG

ATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCATCATGCTGCTATTCCTGCAAATACTGA

AGAAGCATGGGATTTAAATATTTTACTTCTAAATAAATGAATTACTCAATCTCCTATGACCATCTATACATACTC

CACCTTCAAAAAGTACATCAATATTATATCATTAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTT

TTGGACAATGCAATTGTGGCACTGGCACTTATTTCAGTGAAGAAAACTTTGTGGTTCTATGGCATTCATCATTT

GACAAATGCAAGCATCTTCCTTATCAATCAGCTCCTATTGAACTTACTAGCACTGACTGTGGAATCCTTAAGGGC

CCATTACATTTCTGAAGAAGAAAGCTAAG**ATG**AAGGACATGCCACTCCGAATTCATGTGCTACTTGGCCTAGCTA

TCACTACACTAGTACAAGCTGTAGATAAAAAAGTGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGT

TTACACCCAGATCCATTTATATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCCAGCCA

GATTGCCAGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAGACTTTC

CAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAATATTAATGTAAAAAAGATGC

CTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAAACTTACTGAACTGCCTGAAAAATGTCTGTCCGAACTGAGCA

ACTTACAAGAACTCTATATTAATCACAACTTGCTTTCTACAATTTCACCTGGAGCCTTTATTGGCCTACATAATC

TTCTTCGACTTCATCTCAATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAG

AGATTCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACTTTAAGCCTCTTATCAATCTTCGCA

GCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTTGGACTGGAAAACTTAGAAAGCA

TCTCTTTTTACGATAACAGGCTTATTAAAGTACCCCATGTTGCTCTTCAAAAAGTTGTAAATCTCAAATTTTTGG

ATCTAAATAAAAATCCTATTAATAGAATACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAAGAGTTGGGGA

TAAATAATATGCCTGAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAG

CTACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGGAATCACTCATGC

TGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTGCCAAACCTCAAGGAAATCAGCATAC

ACAGTAACCCCATCAGGTGTGACTGTGTCATCCGTTGGATGAACATGAACAAAACCAACATTCGATTCATGGAGC

CAGATTCACTGTTTTGCGTGGACCCACCTGAATTCCAAGGTCAGAATGTTCGGCAAGTGCATTTCAGGGACATGA

TGGAAATTTGTCTCCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGTTT

CCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTGGTCAAAAACTCTTGC

CTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTAGATATAAATGGCGTAACTCCCAAAGAAG

GGGGTTTATATACTTGTATAGCAACTAACCTAGTTGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGAT

CTTTTCCACAAGATAACAATGGCTCTTTGAATATTAAAATAAGAGATATTCAGGCCAATTCAGTTTTGGTGTCCT

GGAAAGCAAGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCATGCTG

CGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATCCATCAACTGAGTATAAAA

TTTGTATTGATATTCCCACCATCTATCAGAAAAACAGAAAAAAATGTGTAAATGTCACCACCAAAGGTTTGCACC

CTGATCAAAAGAGTATGAAAAGAATAATACCACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTG

GTGTGATATGTCTTATCAGCTGCCTCTCTCCAGAAATGAACTGTGATGGTGGACACAGCTATGTGAGGAATTACT

TACAGAAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAAAGAAAAA

GTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCC**TAA**AAACCACCAAGGAAACCTA

CTCCAAAAATGAAC

FIGURE 298

MKDMPLRIHVLLGLAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLTF
PARLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLEEN
KLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDALPN
LEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRLIKVP
HVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDSLAVDNLPDLRKIE
ATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSNPIRCDCVIRW
MNMNKTNIRFMEPDSLFCVDPPEFQGQNVRQVHFRDMMEICLPLIAPESFPSNLNVEAGSYVS
FHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTCIATNLVGA
DLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFVKTENSHAAQ
SARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKEYEKNNTTTLM
ACLGGLLGIIGVICLISCLSPEMNCDGGHSYVRNYLQKPTFALGELYPPLINLWEAGKEKSTS
LKVKATVIGLPTNMS

Important features:

Signal sequence:
amino acids 1-22

Transmembrane domain:
amino acids 633-650

N-glycosylation site.
amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,
608-612, 624-628, 625-629

Casein kinase II phosphorylation site.
amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.
amino acids 570-579

N-myristoylation site.
amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,
491-497, 492-498, 634-640, 702-708

Cell attachment sequence.
amino acids 277-280

FIGURE 299

GCTGTGGGAACCTCTCCACGCGCACGAACTCAGCCAACGATTTCTGATAGATTTTTGGGAGTT

TGACCAGAGATGCAAGGGGTGAAGGAGCGCTTCCTACCGTTAGGGAACTCTGGGGACAGAGCG

CCCCGGCCGCCTGATGGCCGAGGCAGGGTGCGACCCAGGACCCAGGACGGCGTCGGGAACCAT

ACC**ATG**GCCCGGATCCCCAAGACCCTAAAGTTCGTCGTCGTCATCGTCGCGGTCCTGCTGCCA

GTCCTAGCTTACTCTGCCACCACTGCCCGGCAGGAGGAAGTTCCCCAGCAGACAGTGGCCCCA

CAGCAACAGAGGCACAGCTTCAAGGGGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAACAT

ACTGGAGCCTGTAACCCGTGCACAGAGGGTGTGGATTACACCAACGCTTCCAACAATGAACCT

TCTTGCTTCCCATGTACAGTTTGTAAATCAGATCAAAAACATAAAAGTTCCTGCACCATGACC

AGAGACACAGTGTGTCAGTGTAAAGAAGGCACCTTCCGGAATGAAAACTCCCCAGAGATGTGC

CGGAAGTGTAGCAGGTGCCCTAGTGGGGAAGTCCAAGTCAGTAATTGTACGTCCTGGGATGAT

ATCCAGTGTGTTGAAGAATTTGGTGCCAATGCCACTGTGGAAACCCCAGCTGCTGAAGAGACA

ATGAACACCAGCCCGGGGACTCCTGCCCCAGCTGCTGAAGAGACAATGAACACCAGCCCAGGG

ACTCCTGCCCCAGCTGCTGAAGAGACAATGACCACCAGCCCGGGGACTCCTGCCCCAGCTGCT

GAAGAGACAATGACCACCAGCCCGGGGACTCCTGCCCCAGCTGCTGAAGAGACAATGACCACC

AGCCCGGGGACTCCTGCCTCTTCTCATTACCTCTCATGCACCATCGTAGGGATCATAGTTCTA

ATTGTGCTTCTGATTGTGTTTGTT**TGA**AAGACTTCACTGTGGAAGAAATTCCTTCCTTACCTG

AAAGGTTCAGGTAGGCGCTGGCTGAGGGCGGGGGGCGCTGGACACTCTCTGCCCTGCCTCCCT

CTGCTGTGTTCCCACAGACAGAAACGCCTGC

FIGURE 300

MARIPKTLKFVVVIVAVLLPVLAYSATTARQEEVPQQTVAPQQQRHSFKGEECPAGSHRSEHT
GACNPCTEGVDYTNASNNEPSCFPCTVCKSDQKHKSSCTMTRDTVCQCKEGTFRNENSPEMCR
KCSRCPSGEVQVSNCTSWDDIQCVEEFGANATVETPAAEETMNTSPGTPAPAAEETMNTSPGT
PAPAAEETMTTSPGTPAPAAEETMTTSPGTPAPAAEETMTTSPGTPASSHYLSCTIVGIIVLI
VLLIVFV

Important features:

Signal peptide:

Amino acids 1-29

Transmembrane domain:

Amino acids 240-259

N-glycosylation site:

Amino acids 77-81;140-144;156-160

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 126-130

N-myristoylation sites:

Amino acids 56-62;72-78;114-120;154-160;233-239

FIGURE 301

CACAAGCATCTTAATTTGAATCCACAAAGTTTCATGTAATGAAAAGAAATACATAATTTTAAT

TCAACCCGAGTGTTTTCCAAGAAGATTGTATTTGCTTAAATTGCTACAGTAATTCAAGAGACA

GCCCTGTCTGGACACAGAGTTACTGTGGATTTTTAAGAGACTCAGTTAAAGAATTTAGGAATT

TCTGATTCATTTAAAGGATTTACAAATTCATCAACCCCTGAAAACTAAAGCAAATTGAACAGG

AAAAAAAAAAAGAAG**ATG**GGTTTTTTAAGTCCAATATATGTTATTTTCTTCTTTTTTGGAGTC

AAAGTACATTGCCAATATGAAACTTATCAGTGGGATGAAGACTATGACCAAGAGCCAGATGAT

GATTACCAAACAGGATTCCCATTTCGTCAAAATGTAGACTACGGAGTTCCTTTTCATCAGTAT

ACTTTAGGCTGTGTCAGTGAATGCTTCTGTCCAACTAACTTTCCATCATCAATGTACTGTGAT

AATCGCAAACTCAAGACTATCCCAAATATTCCGATGCACATTCAGCAACTCTACCTTCAGTTC

AATGAAATTGAGGCTGTGACTGCAAATTCATTCATCAATGCAACTCATCTTAAAGAAATTAAC

CTCAGCCACAACAAAATTAAATCTCAAAAGATTGATTATGGTGTGTTTGCTAAGCTTCCAAAT

CTACTACAACTTCATCTAGAGCATAATAATTTAGAAGAATTTCCATTTCCTCTTCCTAAATCT

CTGGAAAGACTCCTTCTTGGTTACAATGAAATCTCCAAACTGCAGACAAATGCTATGGATGGG

CTAGTAAACTTGACCATGCTTGATCTCTGTTATAATTATCTTCATGATTCTCTGCTAAAAGAC

AAAATCTTTGCCAAAATGGAAAAACTAATGCAGCTCAACCTCTGCAGTAACAGATTAGAATCA

ATGCCTCCTGGTTTGCCTTCTTCACTTATGTATCTGTCTTTAGAAAATAATTCAATTTCTTCT

ATACCCGAAAAATACTTCGACAAACTTCCAAAACTTCATACTCTAAGAATGTCACACAACAAA

CTACAAGACATCCCATATAATATTTTTAATCTTCCCAACATTGTAGAACTCAGTGTTGGACAC

AACAAATTGAAGCAAGCATTCTATATTCCAAGAAATTTGGAACACCTATACCTACAAAATAAT

GAAATAGAAAGATGAATCTTACAGTGATGTGTCCTTCTATTGACCCACTACATTACCACCAT

TTAACATACATTCGTGTGGACCAAAATAAACTAAAGAACCAATAAGCTCATACATCTTCTTC

TGCTTCCCTCATATACACTATTTATTATGGTGAACAACGAAGCACTAATGGTCAAACAATA

CAACTAAAGACACAAGTTTTCAGGAGATTTCCAGATGATGATGATGAAAGTGAAGATCACGAT

GATCCTGACAATGCTCATGAGAGCCCAGAACAAGAAGGAGCAGAAGGGCACTTTGACCTTCAT

TATTATGAAAATCAAGAA**TAG**CAAGAAACTATATAGGTATACACTTACGACTTCACAAAACCTA

TACTTAATATAGTAAATCTAAGTAAACATGTATTACTCAAAGTAATATATTTAGAATTATGTA

TTAGTATAAGATCAGAATTGAATTTAAGTTGTTGGTGACATCTGCATCATTTCATAGGATTAG

AACTTACTCAAAATAATGTAAATCTTTAAAAATATAAATTAGAATGACAAGTGGGAATCATAA

ATTAAACGTTAATGGTTTCTTATGCTCTTTTTAAATATAGAAATATCATGTTAAAGAAAAAAA

AAAAAAA

FIGURE 302

MGFLSPIYVIFFFFGVKVHCQYETYQWDEDYDQEPDDDYQTGFPFRQNVDYGVPFHQYTLGCV

SECFCPTNFPSSMYCDNRKLKTIPNIPMHIQQLYLQFNEIEAVTANSFINATHLKEINLSHNK

IKSQKIDYGVFAKLPNLLQLHLEHNNLEEFPFPLPKSLERLLLGYNEISKLQTNAMDGLVNLT

MLDLCYNYLHDSLLKDKIFAKMEKLMQLNLCSNRLESMPPGLPSSLMYLSLENNSISSIPEKY

FDKLPKLHTLRMSHNKLQDIPYNIFNLPNIVELSVGHNKLKQAFYIPRNLEHLYLQNNEIEKM

NLTVMCPSIDPLHYHHLTYIRVDQNKLKEPISSYIFFCFPHIHTIYYGEQRSTNGQTIQLKTQ

VFRRFPDDDDESEDHDDPDNAHESPEQEGAEGHFDLHYYENQE

Important feastures:

N-glycosylation sites:

Amino acids 113-117;121-125; 187-191;242-246;316-320

Tyrosine kinase phosphorylation sites:

Amino acids 268-275;300-307

N-myristoylation site:

Amino acids 230-236

Leucine zipper patterns:

Amino acids 146-168;217-239

FIGURE 303

GCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAGC

TGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTACCA

CGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGC**ATG**AATCTGGT

AGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTATGAT

ACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGGGGTTT

AAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGAAACAGT

CTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGGACCTCCA

TCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCATGCCTTCAA

AGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAGTGTGCACAA

AAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGCACTGCGACTG

TACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCACAACGTGATCTG

TAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGCCAACGACGCTGA

CCTTTGTAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCATGTTTGGCTGGTT

CACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAGGATGCCCGGAGACA

CCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGAACCTGATGATATTAG

CACTGTGGTA**TAG**TGTCCAAACTGACTGTCATTGAGAAAGAAAGAAAGTAGTTTGCGATTGCA

GTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAAACTTTGTATTTCAGTTT

TTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACATAAATAATTTGAGTTTAG

GTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGTAAGCTACTATCTGAACATT

AGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTTAATTTAAAAGCAAATAAAAG

CTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAAAACA

FIGURE 304

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDLP
PETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDLSDNRIQ
SVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFLNAA
NDADLCNLPKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKKADEP
DDISTVV

Important features:

Signal sequence:

Amino acids 1-33

Transmembrane domain:

Amino acids 204-219

N-glycosylation sites:

Amino acids 47-51;94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation sites:

Amino acids 37-43;45-51;110-116

FIGURE 305

CGCCACCACTGCGGCCACCGCCA**ATG**AAACGCCTCCCGCTCCTAGTGGTTTTTTTCCACTTTGTTGAATTGTTCCT
ATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGAAATACGCAATGGAATTGAAGCCTGCT
ATTGCAACATGGGATTTTCAGGAAATGGTGTCACAATTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGT
CCTGTGGCGAAAATGCTAATTGCACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCA
GCAGTAACCAAGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATTTAG
ATAATGTCTGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAAGAACCTGTGGCTTTGC
TACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGATATAATTACATATATAGAAATATTAGCTG
AATCATCTTCATTACTAGGTTACAAGAACAACACTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTG
AATTTGTAAAAACCGTGAATAATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATA
GGAGAACACATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCCAAAAGA
CCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTTGATTCATATAACATGAAACATA
TTCATCCTCATATGAATATGGATGGAGACTACATAAATATATTTCCAAAGAGAAAAGCTGCATATGATTCAAATG
GCAATGTTGCAGTTGCATTTTTATATTATAAGAGTATTGGTCCTTTGCTTTCATCATCTGACAACTTCTTATTGA
AACCTCAAAATTATGATAATTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAA
ACCCACCCACATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATAGGTATAGGA
GTCTATGTGCATTTTGGAATTACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAGGGCTGTGAGCTGACAT
ACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACATTTTGCAATTTTGATGTCCTCTGGTCCTT
CCATTGGTATTAAAGATTATAATATTCTTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTG
CCATATGCATTTTTACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCT
GTAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCTTCTGTTCAATCA
TTGCCGGACTGCTACACTACTTCTTTTTTAGCTGCTTTTGCATGGATGTGCATTGAAGGCATACATCTCTATCTCA
TTGTTGTGGGTGTCATCTACAACAAGGGATTTTTGCACAAGAATTTTTATATCTTTGGCTATCTAAGCCCAGCCG
TGGTAGTTGGATTTTCGGCAGCACTAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAA
ACAACTTTATTTGGAGTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGCTTTTGGAGTCATCA
TATACAAAGTTTTTCGTCACACTGCAGGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACATAAGGTCTTGTGCAA
GAGGAGCCCTCGCTCTTCTGTTCCTTCTCGGCACCACCTGGATCTTTGGGGTTCTCCATGTTGTGCACGCATCAG
TGGTTACAGCTTACCTCTTCACAGTCAGCAATGCTTTCCAGGGGATGTTCATTTTTTTATTCCTGTGTGTTTTAT
CTAGAAAGATTCAAGAAGAATATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAGG**TAA**ACAT
AGAGAATGGTGGATAATTACAACTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAATGACT
CATCAAATTATCCAATTATTAACTACTAGACAAAAGTATTTTAAATCAGTTTTTCTGTTTATGCTATAGGAACT
GTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGTTTTTCTATGTGAAATAGTTCTGTCAAAAATA
GTATTGCAGATATTTGGAAAGTAATTGGTTTCTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTTCTA
ACACGAGAAGTATATGAATGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTTGCCTTTGAAACT
AGTCCCCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGAATATCAAA
CAGTGAAAAGGGAATGATAAGATGTATTTTGAATGAACTGTTTTTTCTGTAGACTAGCTGAGAAATTGTTGACAT
AAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAATTGTTCTGAACTTAAATGTCCACTAAAACAACTT
AGACTTCTGTTTGCTAAATCTGTTTCTTTTTCTAATATTCTAAAAAAAAAAAAAAAGGTTTACCTCCACAAATTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 306

```
MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNECGNLTQSCGENANC
TNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAANINKTLTKIRSIKEPVALLQEVYRNS
VTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTLSNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKL
MHTVEQATLRISQSFQKTTEFDTNSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAFL
YYKSIGPLLSSSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAFWNY
SPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLGIIISLICLAICIFTFW
FFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSIIAGLLHYFFLAAFAWMCIEGIHLYLIVVGVIYN
KGFLHKNFYIFGYLSPAVVVGFSAALGYRYYGTTKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHT
AGLKPEVSCFENIRSCARGALALLFLLGTTWIFGVLHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEY
YRLFKNVPCCFGCLR
```

Important features:

Signal peptide:

Amino acids 1-19

Transmembrane domain:

Amino acids 431-450;494-515;573-594;619-636;646-664

N-glycosylation sites:

Amino acids 15-19;21-25;64-68;74-78;127-131;177-181;
188-192;249-253;381-385;395-399

Glycosaminoglycan attachment site:

Amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 360-364

Tyrosine kinase phosphorylation sites:

Amino acids 36-44;670-677

N-myristoylation sites:

Amino acids 38-44;50-56;52-58;80-86;382-388;388-394;
434-440;480-486;521-527

Aspartic acid and asparagine hydroxylation site:

Amino acids 75-87

FIGURE 307

CCAGGCCGGGAGGCGACGCGCCCAGCCGTCTAAACGGGAACAGCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGT
ATCTGACGGCGCCAGGTTGCGTAGGTGCGGCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGC**ATG**GC
CCGGAGGAGCGCCTTCCCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGA
GGCCGGGCCGCCGCAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCATAGGATTTGA
AGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGATTTCAGAAAAGCGCAACAGAGAAT
GCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTACCTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGA
ATTCCTGTCCTTGCGCTCCCTGGATAAAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGT
GCCTCACAAGGCATCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGT
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCTTCTTTAAAACATG
TCAACAAGCTGAGTGCCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAATGAAAGACGCATCTGCGAGTGTCCTGA
TGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTGTACCCCACGATGTATGAATGGTGGACTTTGTGTGACTCC
TGGTTTCTGCATCTGCCCACCTGGATTCTATGGAGTGAACTGTGACAAAGCAAACTGCTCAACCACCTGCTTTAA
TGGAGGGACCTGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAGCAA
ATGCCCACAACCCTGTCGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTTCCAAAGGTTACCAGGG
AGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACATGGAACCTGCCATGAACCCAACAAATGCCA
ATGTCAAGAAGGTTGGCATGGAAGACACTGCAATAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGC
AGGCGCCCAGCTCAGGCAGCACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTA
CATCTGG**TGA**ACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCATGTGTT
GAATGTTCAAATAATGTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCTTCATTATAAATCACTGAGC
TGATATTTACTCTTCCTTTTAAGTTTTCTAAGTACGTCTGTAGCATGATGGTATAGATTTTCTTGTTTCAGTGCT
TTGGGACAGATTTTATATTATGTCAATTGATCAGGTTAAAATTTTCAGTGTGTAGTTGGCAGATATTTTCAAAAT
TACAATGCATTTATGGTGTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAA
GAATTTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGATGTTTGTTAC
ATTTTTAAAAATTGCTCTTAATTTTTAAACTCTCAATACAATATATTTTGACCTTACCATTATTCCAGAGATTCA
GTATTAAAAAAAAAAAAATTACACTGTGGTAGTGGCATTTAAACAATATAATATATTCTAAACACAATGAAATAG
GGAATATAATGTATGAACTTTTTGCATTGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCT
AATAAACATTTTATACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAA
AAAAAAAA

FIGURE 308

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGKM
APFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLSLRSLDKGIMADPTVNVPLLGT
VPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGCRNG
GFCNERRICECPDGFHGPHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCSTTCFN
GGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVCEPGCG
AHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDPPESNYIW

Important features:

Signal sequence:
Amino acids 1-28

N-glycosylation sites:
Amino acids 88-92;245-249

Tyrosine kinase phosphorylation site:
Amino acids 370-378

N-myristoylation sites:
Amino acids 184-190;185-191;189-195;315-321

ATP/GTP-binding site motif A (P-loop):
Amino acids 285-293

EGF-like domain cysteine pattern signatures:
Amino acids 198-210;230-242;262-274;294-306;326-338

FIGURE 309

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGGG
TTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGACT
CTTGCAAGCTGGATGCCCTCTGTGGATGAAAG**ATG**TATCATGGAATGAACCCGAGCAATGGAG
ATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCCAGAGACTCTTGGCCG
TGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTTCGATG
ACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCGGAGGGG
TTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAGGGCGCTA
CAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGATAATTCCA
TCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGÄGATTCATAACAAGACAT
ATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGGTACCCCGACC
TACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCCCATCTGTCAAG
GCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGCTCCAGACCTCCT
TCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTTGATGGGTCTGCGT
ATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTGCCTTGCTCTGGAAG
CCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCGTCTGCCACCCGCGGCCTT
GTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCTGGCTACAGCCTCACCA
GCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTCTTATCAAGTCTACTGCA
TCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGACCACGTGGAAGATTGTGG
CGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATCCTGGCCAGGATGTTCCAGA
CCAAGTTCAAGGCCCACTTTCCCCCCAGGGGGCCTCCCCGGAGTTCCAGCAGTGACCCTGACT
TTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATGACGAAGCTGTGAGTGGCGGCT
TGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAGGGCTGCCCCTTACCCGTGGACG
ACCAGAGCCCCCCAGCATACCCCGGCTCAGGGGACACGGACACAGGCCCAGGGGAGTCAGAAA
CCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAAGTCTGTATTCACCTCCCAGGTGCC
AAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATAATTGCCAGCACGGCAGAGGAGGTGG
CATCCACCAGCCCAGGCATCCATCATGCCCACTGGGTGTTGTTCCTAAGAAAC**TGA**TTGATTA
AAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAAATACATGTTGATCTGTGGAGTTGATTC
CTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAAAGCTCTGATCCTTAAAATTGCTATGCTG
ATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTCCTGTTTCTTCTTGACACAGACTGATTAAA
AATTAAAAGNAAAAAA

FIGURE 310

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGIP
ENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRIPQ
IEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLASSNG
YVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLPPMVS
HGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQTWPSTH
ETLLTTWKIVAFTATSVLLVLLLVILARMFQTKFKAHFPPRGPPRSSSSDPDFVVVDGVPVML
PSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGDTDTGPGESETCDSVSGSSEL
LQSLYSPPRCQESTHPASDNPDIIASTAEEVASTSPGIHHAHWVLFLRN

Important features:
Signal sequence:
amino acids 1-41

Transmembrane domain:
amino acids 325-344

N-glycosylation site.
amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.
amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366, 364-368, 380-384, 467-471, 468-472

N-myristoylation site.
amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424, 478-484

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 394-405

FIGURE 311

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGC**ATG**AGCGGCGGTTGGATGGCGCAGGTTGGAG

CGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCTGG

AGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAGGCT

CGTGCCCACCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCTCACCTGGCGCT

GCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCCATGTA

CCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCAGTGACT

GCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCAGGCGAGC

TCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCACCCAGACT

GTCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAGGGGATGCCA

CAACCATGGGGCCCCTGTGACCCTGGAGAGTGTCACCTCTCTCAGGAATGCCACAACCATGG

GGCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTCGGGAATGCCACATCCTCCTCTGCCGGAG

ACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGCTCAGTGCAAGCC

TGGTCACCGCCACCCTCCTCCTTTTGTCCTGGCTCCGAGCCCAGGAGCGCCTCCGCCCACTGG

GGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGACCTCGCTGCCC**TGA**G

GACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACAGGAGGAGAGCAGTGA

TGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTTCTGGCCACGTGGAACC

TCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCCTGGACACTCCCTATGGA

GATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAGGGGCTGGCCCCAGGCAGC

TCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTGCCCCGTCTGAGGGTGGCGA

TTAAAGTTGCTTC

FIGURE 312

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCRT
SGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGTDKKLRN
CSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVTLES
VTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSASLVTATLLLLSW
LRAQERLRPLGLLVAMKESLLLSEQKTSLP

Important features:

Signal sequence:

Amino acids 1-30

Transmembrane domain:

Amino acids 231-248

N-glycosylation sites:

Amino acids 126-130;195-199;213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation sites:

Amino acids 3-9;10-16;26-32;30-36;112-118;166-172;212-218;
224-230;230-236;263-269

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 44-55

Leucine zipper pattern:

Amino acids 17-39

FIGURE 313

CGGACGCGTGGGCGTCCGGCGGTCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGGCCCCAGCCCACAC
CTTCACCAGGGCCCAGGAGCCACC**ATG**TGGCGATGTCCACTGGGGCTACTGCTGTTGCTGCCGCTGGCTGGCCAC
TTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGGGAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGAC
GCGGGAGGCCGGTACTGCCAGGAGCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTG
GGCGCCATCTGTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTCTGC
CTCGGCGTGCCACCCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTATCCAGTCTTGGGAACG
TACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGTGGCATGGTGGATCCAGACATGATCAAAGC
CATCAACCAGGGCAACTATGGCTGGCAGGCTGGGAACCACAGCGCCTTCTGGGGCATGACCCTGGA**TGA**GGGCAT
TCGCTACCGCCTGGGCACCATCCGCCCATCTTCCTCGGTCATGAACATGCATGAAATTTATACAGTGCTGAACCC
AGGGGAGGTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGACCAAGG
CAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCTCAATCCATTCTCTGGGACA
CATGACGCCTGTCCTGTCGCCCCAGAACCTGCTGTCTTGTGACACCCACCAGCAGCAGGGCTGCCGCGGTGGGCG
TCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGA
ACGAGACGAGGCTGGCCCTGCGCCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCAC
TGCCCACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTCCAA
CGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGGTGCATGAGGACTTCTT
CCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGGAGGCCAGAGAGATACCGCCGGCATGGGAC
CCACTCAGTCAAGATCACAGGATGGGGAGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGC
CAACTCCTGGGGCCCAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGGCGTCAATGAGTGCGACATCGA
GAGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTCATCACTGAGGCTGCGGGCACCACGC
GGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCCAATGGGGCGGTGACCCCAGCCTCGCCCGA
CAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAATCCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCCTGGG
AGCCGCGGGCAGGCGAGACTGGCGGAGCCCCCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAG
CACAGCTGCAGATCCCAGGCCTCTGGCGCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGCCC
CAATACCCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCGTTGCCCAGGTTGGAG
TGCAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCAAGTGACCCTCCCACCTCAGCCTCTCAAG
TAGCTGGGACTACAGGTGCACCACCACACCTGGCTAATTTTTGTATTTTTTGTAAAGAGGGGGGTCTCACTGTGT
TGCCCAGGCTGGTTTCGAACTCCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGG
CATGAGCCACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTTCTTTTCACTGTTTTAAAATAAAA
CCAAAGTATTGATAAAAAAAAA

FIGURE 314

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRADDCA
LPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPFPPIQGCMHGGRIYPVLGTYWDNCNRCT
CQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

Important features:

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300, 411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230, 269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

FIGURE 315

CGGACGCGTGGGCCCCTGGTGGGCCCAGCAAG**ATG**GATCTACTGTGGATCCTGCCCTCCCTGT
GGCTTCTCCTGCTTGGGGGGCCTGCCTGCCTGAAGACCCAGGAACACCCCAGCTGCCCAGGAC
CCAGGGAACTGGAAGCCAGCAAAGTTGTCCTCCTGCCCAGTTGTCCCGGAGCTCCAGGAAGTC
CTGGGGAGAAGGGAGCCCCAGGTCCTCAAGGGCCACCTGGACCACCAGGCAAGATGGGCCCCA
AGGGTGAGCCAGGCCCCAGAAACTGCCGGGAGCTGTTGAGCCAGGGCGCCACCTTGAGCGGCT
GGTACCATCTGTGCCTACCTGAGGGCAGGGCCCTCCCAGTCTTTTGTGACATGGACACCGAGG
GGGGCGGCTGGCTGGTGTTTCAGAGGCGCCAGGATGGTTCTGTGGATTTCTTCCGCTCTTGGT
CCTCCTACAGAGCAGGTTTTGGGAACCAAGAGTCTGAATTCTGGCTGGGAAATGAGAATTTGC
ACCAGCTTACTCTCCAGGGTAACTGGGAGCTGCGGGTAGAGCTGGAAGACTTTAATGGTAACC
GTACTTTCGCCCACTATGCGACCTTCCGCCTCCTCGGTGAGGTAGACCACTACCAGCTGGCAC
TGGGCAAGTTCTCAGAGGGCACTGCAGGGGATTCCCTGAGCCTCCACAGTGGGAGGCCCTTTA
CCACCTATGACGCTGACCACGATTCAAGCAACAGCAACTGTGCAGTGATTGTCCACGGTGCCT
GGTGGTATGCATCCTGTTACCGATCAAATCTCAATGGTCGCTATGCAGTGTCTGAGGCTGCCG
CCCACAAATATGGCATTGACTGGGCCTCAGGCCGTGGTGTGGGCCACCCCTACCGCAGGGTTC
GGATGATGCTTCGA**TAG**GGCACTCTGGCAGCCAGTGCCCTTATCTCTCCTGTACAGCTTCCGG
ATCGTCAGCCACCTTGCCTTTGCCAACCACCTCTGCTTGCCTGTCCACATTTAAAAATAAAAT
CATTTTAGCCCTTTCA

FIGURE 316

MDLLWILPSLWLLLLGGPACLKTQEHPSCPGPRELEASKVVLLPSCPGAPGSPGEKGAPGPQG
PPGPPGKMGPKGEPGPRNCRELLSQGATLSGWYHLCLPEGRALPVFCDMDTEGGGWLVFQRRQ
DGSVDFFRSWSSYRAGFGNQESEFWLGNENLHQLTLQGNWELRVELEDFNGNRTFAHYATFRL
LGEVDHYQLALGKFSEGTAGDSLSLHSGRPFTTYDADHDSSNSNCAVIVHGAWWYASCYRSNL
NGRYAVSEAAAHKYGIDWASGRGVGHPYRRVRMMLR

Important features:

Signal peptide:

Amino acids 1-16

N-glycosylation site:

Amino acids 178-182

Glycosaminoglycan attachment site:

Amino acids 272-276

Tyrosine kinase phosphorylation site:

Amino acids 188-197

N-myristoylation sites:

Amino acids 16-22;89-95;144-150;267-273

Fibrinogen beta and gamma chains C-terminal domain signature:

Amino acids 242-255

FIGURE 317

CCCAAGCCAGCCGAGCCGCCAGAGCCGCGGGCCGCGGGGGTGTCGCGGGCCCAACCCCAGG**AT**
GCTCCCCTGCGCCTCCTGCCTACCCGGGTCTCTACTGCTCTGGGCGCTGCTACTGTTGCTCTT
GGGATCAGCTTCTCCTCAGGATTCTGAAGAGCCCGACAGCTACACGGAATGCACAGATGGCTA
TGAGTGGGACCCAGACAGCCAGCACTGCCGGGATGTCAACGAGTGTCTGACCATCCCTGAGGC
CTGCAAGGGGGAAATGAAGTGCATCAACCACTACGGGGGCTACTTGTGCCTGCCCCGCTCCGC
TGCCGTCATCAACGACCTACATGGCGAGGGACCCCCGCCACCAGTGCCTCCCGCTCAACACCC
CAACCCCTGCCCACCAGGCTATGAGCCCGACGATCAGGACAGCTGTGTGGATGTGGACGAGTG
TGCCCAGGCCCTGCACGACTGTCGCCCCAGCCAGGACTGCCATAACTTGCCTGGCTCCTATCA
GTGCACCTGCCCTGATGGTTACCGCAAGATCGGGCCCGAGTGTGTGGACATAGACGAGTGCCG
CTACCGCTACTGCCAGCACCGCTGCGTGAACCTGCCTGGCTCCTTCCGCTGCCAGTGCGAGCC
GGGCTTCCAGCTGGGGCCTAACAACCGCTCCTGTGTTGATGTGAACGAGTGTGACATGGGGGC
CCCATGCGAGCAGCGCTGCTTCAACTCCATGGGACCTTCCTGTGTCGCTGCCACCAGGGCTA
TGAGCTGCATCGGGATGGCTTCTCCTGCAGTGATATTGATGAGTGTAGCTACTCCAGCTACCT
CTGTCAGTACCGCTGCGTCAACGAGCCAGGCCGTTTCTCCTGCCACTGCCCACAGGGTTACCA
GCTGCTGGCCACACGCCTCTGCCAAGACATTGATGAGTGTGAGTCTGGTGCGCACCAGTGCTC
CGAGGCCCAAACCTGTGTCAACTTCCATGGGGGCTACCGCTGCGTGGACACCAACCGCTGCGT
GGAGCCCTACATCCAGGTCTCTGAGAACCGCTGTCTCTGCCCGGCCTCCAACCCTCTATGTCG
AGAGCAGCCTTCATCCATTGTGCACCGCTACATGACCATCACCTCGGAGCGGAGCGTGCCCGC
TGACGTGTTCCAGATCCAGGCGACCTCCGTCTACCCCGGTGCCTACAATGCCTTTCAGATCCG
TGCTGGAAACTCGCAGGGGGACTTTTACATTAGGCAAATCAACAACGTCAGCGCCATGCTGGT
CCTCGCCCGGCCGGTGACGGGCCCCGGGAGTACGTGCTGGACCTGGAGATGGTCACCATGAA
TTCCCTCATGAGCTACCGGGCCAGCTCTGTACTGAGGCTCACCGTCTTTGTAGGGGCCTACAC
CTTC**TGA**GGAGCAGGAGGGAGCCACCCTCCCTGCAGCTACCCTAGCTGAGGAGCCTGTTGTGA
GGGGCAGAATGAGAAAGGCAATAAAGGGAGAAAGAAAGTCCTGGTGGCTGAGGTGGGCGGGTC
ACACTGCAGGAAGCCTCAGGCTGGGGCAGGGTGGCACTTGGGGGGGCAGGCCAAGTTCACCTA
AATGGGGGTCTCTATATGTTCAGGCCCAGGGGCCCCCATTGACAGGAGCTGGGAGCTCTGCAC
CACGAGCTTCAGTCACCCCGAGAGGAGAGGAGGTAACGAGGAGGGCGGACTCCAGGCCCCGGC
CCAGAGATTTGGACTTGGCTGGCTTGCAGGGGTCCTAAGAAACTCCACTCTGGACAGCGCCAG
GAGGCCCTGGGTTCCATTCCTAACTCTGCCTCAAACTGTACATTTGGATAAGCCCTAGTAGTT
CCCTGGGCCTGTTTTTCTATAAAACGAGGCAACTGGAAAAAAAAAAA

FIGURE 318

MLPCASCLPGSLLLWALLLLLLGSASPQDSEEPDSYTECTDGYEWDPDSQHCRDVNECLTIPE
ACKGEMKCINHYGGYLCLPRSAAVINDLHGEGPPPPVPPAQHPNPCPPGYEPDDQDSCVDVDE
CAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCQHRCVNLPGSFRCQCE
PGFQLGPNNRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGFSCSDIDECSYSSY
LCQYRCVNEPGRFSCHCPQGYQLLATRLCQDIDECESGAHQCSEAQTCVNFHGGYRCVDTNRC
VEPYIQVSENRCLCPASNPLCREQPSSIVHRYMTITSERSVPADVFQIQATSVYPGAYNAFQI
RAGNSQGDFYIRQINNVSAMLVLARPVTGPREYVLDLEMVTMNSLMSYRASSVLRLTVFVGAYTF

Important features:

Signal sequence:

Amino acids 1-25

N-glycosylation sites:

Amino acids 198-202;394-398

N-myristoylation sites:

Amino acids 76-82;145-151;182-188;222-228;290-296;305-311;
371-377;381-387

Aspartic acid and asparagine hydroxylation sites:

amino acids 140-152;177-189;217-229;258-270

FIGURE 319

GCTGGGGAC**ATG**AGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTC
TCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCTGGCCACCTCCCCGATGC
CCGCTGGGAGTACCCCTGGTGCTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGGCTG
GGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGCCAGCCCGGG
GCAGGACCCGGTGGCCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGTGAGGTG
AACGGCCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGC
GAGGACGGCGGCTTCACCTGCGTGCCGCTGTGCAGCGAGGATGTGCGGCTGCCCAGCTGGGAC
TGCCCCCACCCCAGGAGGGTCGAGGTCCTGGGCAAGTGCTGCCCTGAGTGGGTGTGCGGCCAA
GGAGGGGGACTGGGGACCCAGCCCCTTCCAGCCCAAGGACCCCAGTTTTCTGGCCTTGTCTCT
TCCCTGCCCCCTGGTGTCCCCTGCCCAGAATGGAGCACGGCCTGGGGACCCTGCTCGACCACC
TGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAGAACCGCTTCTGCCGACTGGAGACCCAG
CGCCGCCTGTGCCTGTCCAGGCCCTGCCCACCCTCCAGGGGTCGCAGTCCACAAAACAGTGCC
TTC**TAG**AGCCGGGCTGGGAATGGGGACACGGTGTCCACCATCCCCAGCTGGTGGCCCTGTGCC
TGGGCCCTGGGCTGATGGAAGATGGTCCGTGCCCAGGCCCTTGGCTGCAGGCAACACTTTAGC
TTGGGTCCACCATGCAGAACACCAATATTAACACGCTGCCTGGTCTGTCTGGATCCCGAGGTA
TGGCAGAGGTGCAAGACCTAGTCCCCTTTCCTCTAACTCACTGCCTAGGAGGCTGGCCAAGGT
GTCCAGGGTCCTCTAGCCCACTCCCTGCCTACACACACAGCCTATATCAAACATGCACACGGG
CGAGCTTTCTCTCCGACTTCCCCTGGGCAAGAGATGGGACAAGCAGTCCCTTAATATTGAGGC
TGCAGCAGGTGCTGGGCTGGACTGGCCATTTTTCTGGGGGTAGGATGAAGAGAAGGCACACAG
AGATTCTGGATCTCCTGCTGCCTTTTCTGGAGTTTGTAAAATTGTTCCTGAATACAAGCCTAT
GCGTGA

FIGURE 320

MRGTPKTHLLAFSLLCLLSKVRTQLCPTPCTCPWPPPRCPLGVPLVLDGCGCCRVCARRLGEP
CDQLHVCDASQGLVCQPGAGPGGRGALCLLAEDDSSCEVNGRLYREGETFQPHCSIRCRCEDG
GFTCVPLCSEDVRLPSWDCPHPRRVEVLGKCCPEWVCGQGGGLGTQPLPAQGPQFSGLVSSLP
PGVPCPEWSTAWGPCSTTCGLGMATRVSNQNRFCRLETQRRLCLSRPCPPSRGRSPQNSAF

Important features:

Signal sequence:
Amino acids 1-23

N-myristoylation sites:
Amino acids 3-9;49-55;81-87;85-91;126-132;164-170;166-172;
167-173;183-189;209-215

Insulin-like growth factor binding proteins signature:
Amino acids 49-65

von Willebrand C1 domain:
Amino acids 107-124

Thrombospondin 1 Homology Block:
Amino acids 201-216

IGF binding protein site:
Amino acids 49-58

FIGURE 321

AGAACCTCAGAAATGTGAGTTATTTGGGAATGGCTGTTTGTAAATGTCCTTACGTAAGCCAAG
AGGAGGTCTTGACTTGGGGTCCCAGGGGTACCGCAGATCCCAGGGACTGGAGCAGCACTAGCA
AGCTCTGGAGGATGAGCCAGGAGTCTGGAATTGAGGCTGAGCCAAAGACCCCAGGGCCGTCTC
AGTCTCATAAAAGGGGATCAGGCAGGAGGAGTTTGGGAGAAACCTGAGAAGGGCCTGATTTGC
AGCATC**ATG**ATGGGCCTCTCCTTGGCCTCTGCTGTGCTCCTGGCCTCCCTCCTGAGTCTCCAC
CTTGGAACTGCCACACGTGGGAGTGACATATCCAAGACCTGCTGCTTCCAATACAGCCACAAG
CCCCTTCCCTGGACCTGGGTGCGAAGCTATGAATTCACCAGTAACAGCTGCTCCCAGCGGGCT
GTGATATTCACTACCAAAAGAGGCAAGAAAGTCTGTACCCATCCAAGGAAAAAATGGGTGCAA
AAATACATTTCTTTACTGAAAACTCCGAAACAATTG**TGA**CTCAGCTGAATTTTCATCCGAGGA
CGCTTGGACCCCGCTCTTGGCTCTGCAGCCCTCTGGGGAGCCTGCGGAATCTTTTCTGAAGGC
TACATGGACCCGCTGGGGAGGAGAGGGTGTTTCCTCCCAGAGTTACTTTAATAAAGGTTGTTC
ATAGAGTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 322

MMGLSLASAVLLASLLSLHLGTATRGSDISKTCCFQYSHKPLPWTWVRSYEFTSNSCSQRAVI
FTTKRGKKVCTHPRKKWVQKYISLLKTPKQL

Important features:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 3-9, 26-32

Amidation site.

amino acids 68-72

Small cytokines (intecrine/chemokine).

amino acids 23-88

FIGURE 323

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAG**ATG**CAGGTGAGCAAGAGGATGCTGGCGGGGGGCGTGAGGAGCA

TGCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCTGGGCTCAGTGCTGTCAGGCTCGGCCA

CGGGCTGCCCGCCCCGCTGCGAGTGCTCCGCCCAGGACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAG

TCCCCGAGGGCATCCCCACCGAGACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACG

AGTTCGCCAGCTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCGCCT

TCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATCCCGCTAGGCGTCTTCA

CTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGATCGTTATCCTACTGGACTACATGTTTCAGG

ACCTGTACAACCTCAAGTCACTGGAGGTTGGCGACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCC

TCAACAGCCTGGAGCAGCTGACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGC

ACGGCCTCATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGCTGTACC

GACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAACTGCCTCTACGGCCTCAACC

TGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCCCTACCTGGCCGTCCGCCACCTAGTCTATCTCC

GCTTCCTCAACCTCTCCTACAACCCCATCAGCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGG

AGATCCAGCTGGTGGGCGGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGC

TCAATGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGGAGACACTCA

TCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGGCGCCGCTGGCGGCTCAACTTCA

ACCGGCAGCAGCCCACGTGCGCCACGCCCGAGTTTGTCCAGGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTAC

TGCCCAACTACTTCACCTGCCGCCGCGCCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCC

ACACGGTGCAGTTTGTGTGCCGGGCCGATGGCGACCCGCCGCCCGCCATCCTCTGGCTCTCACCCCGAAAGCACC

TGGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCTACGCCCAGGTAC

AGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGACTCCATGCCCGCCCACCTGCATGTGC

GCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAAGACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGAG

AGGCCAACAGCACCCGCGCCACTGTGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCT

TCATCTCTTTCCTGGGCGTCGTCCTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACACAA

AGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCGACGCGCCCCGCAAGT

TCAACATGAAGATGATA**TGA**GGCCGGGGCGGGGGGCAGGGACCCCCGGGCGGCCGGGCAGGGGAAGGGGCCTGGT

CGCCACCTGCTCACTCTCCAGTCCTTCCCACCTCCTCCCTACCCTTCTACACACGTTCTCTTTCTCCCTCCCGCC

TCCGTCCCCTGCTGCCCCCCGCCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGG

GGACCCCACCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCAATAAT

TCAATAAAAAAGTTACGAACTTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTTTATGAAAACTTGAAATAA

TAAAAAGAGAAAAAAACTAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 324

MQVSKRMLAGGVRSMPSPLLACWQPILLLVLGSVLSGSATGCPPRCECSAQDRAVLCHRKCFVAVPEGIPTETRL
LDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFNNLFNLRTLGLRSNRLKLIPLGVFTGLSNLTKQDI
SENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRAFSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHL
NINAIRDYSFKRLYRLKVLEISHWPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPIS
TIEGSMLHELLRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSVGNLETLILDSNPLACDC
RLLWVFRRRWRLNFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQVFVDEGHTVQFVCRADG
DPPPAILWLSPRKHLVSAKSNGRLTVFPDGTLEVRYAQVQDNGTYLCIAANAGGNDSMPAHLHVRSYSPDWPHQP
NKTFAFISNQPGEGEANSTRATVPFPFDIKTLIIATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPR
KSDAGISSADAPRKFNMKMI

Important features:
Signal sequence:
amino acids 1-41

Transmembrane domain:
amino acids 556-578

N-glycosylation site.
amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496,
505-509, 526-530, 542-546

Casein kinase II phosphorylation site.
amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

Tyrosine kinase phosphorylation site.
amino acids 590-598

N-myristoylation site.
amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326,
344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 32-43

FIGURE 325

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGAG

GCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGAGG

AGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAGGAG

GAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAACGGAG

AGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGAGTAGG

AAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGGAAAGAC

ACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTGGCTGCTT

TGGCATTTGGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCTGGAGGGAC

AGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGGCAGGGGTCC

CTCGGAGGCCTCCTGGGG**ATG**GGGGCTGCAGCTCGTCTGAGCGCCCCTCGAGCGCTGGTACTC

TGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGGACTGGTGGAGC

TACAAGGATAATCTCCAGGGAAACTTCGTGCCAGGGCCTCCTTTCTGGGGCCTGGTGAATGCA

GCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGAGCTGAAGAGGGTT

CTTTATGACCCCTTTCTGCCCCCATTAAGGCTCAGCACTGGAGGAGAGAAGCTCCGGGGAACC

TTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCTGTGGTCAATGTGTCT

GGAGGTCCCCTCCTTTACAGCCACCGACTCAGTGAACTGCGGCTGCTGTTTGGAGCTCGCGAC

GGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTGAGGTGCAGCTCATTCAC

TTCAACCAGGAACTCTACGGGAATTTCAGCGCTGCCTCCCGCGGCCCCAATGGCCTGGCCATT

CTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCCATTCCTCAGTCGCCTCCTTAACCGC

GACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTCTTCAAGACCTGAGCCTGGAG

CTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGCTCTCTCAGCACCCCGCCCTGC

TCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAATATCACCTCCCTTCAGATGCAC

TCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCTTCCAGAGCCTCAGCGGTAACAGC

CGGCCCCTGCAGCCCTTGGCCCACAGGGCACTGAGGGGCAACAGGGACCCCCGGCACCCCGAG

AGGCGCTGCCGAGGCCCCAACTACCGCCTGCATGTGGATGGTGTCCCCCATGGTCGC**TGA**GAC

TCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTCCCCACAAGGCGAGGGGAGTTACCCCT

AAAACAAAGCTATTAAAGGGACAGAATACTTA

326/550

FIGURE 326

MGAAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLCA
VGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPLLY
SHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSLFVN
VASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSETVTW
ILIDRALNITSLQMHSLRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPERRCRGP
NYRLHVDGVPHGR

Important features:

Signal peptide:

Amino acids 1-23

Transmembrane domain:

Amino acids 177-199

N-glycosylation sites:

Amino acids 118-122;170-174;260-264

Eukaryotic-type carbonic anhydrases proteins:

Amino acids 222-271;128-165;45-93

FIGURE 327

GGACTAATCTGTGGGAGCAGTTTATTCCAGTATCACCCAGGGTGCAGCCACACCAGGACTGTGTTGAAGGGTGTT

TTTTTTCTTTTAAATGTAATACCTCCTCATCTTTTCTTCTTACACAGTGTCTGAGAACATTTACATTATAGATAA

GTAGTACATGGTGGATAACTTCTACTTTTAGGAGGACTACTCTCTTCTGACAGTCCTAGACTGGTCTTCTACACT

AAGACACC**ATG**AAGGAGTATGTGCTCCTATTATTCCTGGCTTTGTGCTCTGCCAAACCCTTCTTTAGCCCTTCAC

ACATCGCACTGAAGAATATGATGCTGAAGGATATGGAAGACACAGATGATGATGATGATGATGATGATGATGATG

ATGATGATGAGGACAACTCTCTTTTTCCAACAAGAGAGCCAAGAAGCCATTTTTTTCCATTTGATCTGTTTCCAA

TGTGTCCATTTGGATGTCAGTGCTATTCACGAGTTGTACATTGCTCAGATTTAGGTTTGACCTCAGTCCCAACCA

ACATTCCATTTGATACTCGAATGCTTGATCTTCAAAACAATAAAATTAAGGAAATCAAAGAAAATGATTTTAAAG

GACTCACTTCACTTTATGGTCTGATCCTGAACAACAACAAGCTAACGAAGATTCACCCAAAAGCCTTTCTAACCA

CAAAGAAGTTGCGAAGGCTGTATCTGTCCCACAATCAACTAAGTGAAATACCACTTAATCTTCCCAAATCATTAG

CAGAACTCAGAATTCATGAAAATAAAGTTAAGAAAATACAAAAGGACACATTCAAAGGAATGAATGCTTTACACG

TTTTGGAAATGAGTGCAAACCCTCTTGATAATAATGGGATAGAGCCAGGGGCATTTGAAGGGGTGACGGTGTTCC

ATATCAGAATTGCAGAAGCAAAACTGACCTCAGTTCCTAAAGGCTTACCACCAACTTTATTGGAGCTTCACTTAG

ATTATAATAAAATTTCAACAGTGGAACTTGAGGATTTTAAACGATACAAAGAACTACAAAGGCTGGGCCTAGGAA

ACAACAAATCACAGATATCGAAAATGGGAGTCTTGCTAACATACCACGTGTGAGAGAAATACATTTGGAAAACA

ATAAACTAAAAAAAATCCCTTCAGGATTACCAGAGTTGAAATACCTCCAGATAATCTTCCTTCATTCTAATTCAA

TTGCAAGAGTGGGAGTAAATGACTTCTGTCCAACAGTGCCAAAGATGAAGAAATCTTTATACAGTGCAATAAGTT

TATTCAACAACCCGGTGAAATACTGGGAAATGCAACCTGCAACATTTCGTTGTGTTTTGAGCAGAATGAGTGTTC

AGCTTGGGAACTTTGGAATG**TAA**TAATTAGTAATTGGTAATGTCCATTTAATATAAGATTCAAAAATCCCTACAT

TTGGAATACTTGAACTCTATTAATAATGGTAGTATTATATATACAAGCAAATATCTATTCTCAAGTGGTAAGTCC

ACTGACTTATTTTATGACAAGAAATTTCAACGGAATTTTGCCAAACTATTGATACATAAGGGGTTGAGAGAAACA

AGCATCTATTGCAGTTTCCTTTTTGCGTACAAATGATCTTACATAAATCTCATGCTTGACCATTCCTTTCTTCAT

AACAAAAAAGTAAGATATTCGGTATTTAACACTTTGTTATCAAGCACATTTTAAAAAGAACTGTACTGTAAATGG

AATGCTTGACTTAGCAAAATTTGTGCTCTTTCATTTGCTGTTAGAAAAACAGAATTAACAAAGACAGTAATGTGA

AGAGTGCATTACACTATTCTTATTCTTTAGTAACTTGGGTAGTACTGTAATATTTTTAATCATCTTAAAGTATGA

TTTGATATAATCTTATTGAAATTACCTTATCATGTCTTAGAGCCCGTCTTTATGTTTAAAACTAATTTCTTAAAA

TAAAGCCTTCAGTAAATGTTCATTACCAACTTGATAAATGCTACTCATAAGAGCTGGTTTGGGGCTATAGCATAT

GCTTTTTTTTTTTTAATTATTACCTGATTTAAAAATCTCTGTAAAAACGTGTAGTGTTTCATAAAATCTGTAACT

CGCATTTTAATGATCCGCTATTATAAGCTTTTAATAGCATGAAAATTGTTAGGCTATATAACATTGCCACTTCAA

CTCTAAGGAATATTTTTGAGATATCCCTTTGGAAGACCTTGCTTGGAAGAGCCTGGACACTAACAATTCTACACC

AAATTGTCTCTTCAAATACGTATGGACTGGATAACTCTGAGAAACACATCTAGTATAACTGAATAAGCAGAGCAT

CAAATTAAACAGACAGAAACCGAAAGCTCTATATAAATGCTCAGAGTTCTTTATGTATTTCTTATTGGCATTCAA

CATATGTAAAATCAGAAACAGGGAAATTTTCATTAAAAATATTGGTTTGAAAT

FIGURE 328

MKEYVLLLFLALCSAKPFFSPSHIALKNMMLKDMEDTDDDDDDDDDDDDEDNSLFPTREPRS

HFFPFDLFPMCPFGCQCYSRVVHCSDLGLTSVPTNIPFDTRMLDLQNNKIKEIKENDFKGLTS

LYGLILNNNKLTKIHPKAFLTTKKLRRLYLSHNQLSEIPLNLPKSLAELRIHENKVKKIQKDT

FKGMNALHVLEMSANPLDNNGIEPGAFEGVTVFHIRIAEAKLTSVPKGLPPTLLELHLDYNKI

STVELEDFKRYKELQRLGLGNNKITDIENGSLANIPRVREIHLENNKLKKIPSGLPELKYLQI

IFLHSNSIARVGVNDFCPTVPKMKKSLYSAISLFNNPVKYWEMQPATFRCVLSRMSVQLGNFGM

Important features:

Signal sequence.

amino acids 1-15

N-glycosylation site.

amino acids 281-285

N-myristoylation sites.

amino acids 129-135, 210-216, 214-220, 237-243, 270-276, 282-288

Leucine zipper pattern.

amino acids 154-176

FIGURE 329

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGCA

TCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGTGT

TTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACC**ATG**ATCACTGGTGTGTT

CAGCATGCGCTTGTGGACCCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACCAGCG

GCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGCCTGCT

GAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCCGCTCCC

GCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACTCAGTTTGA

TTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCTCAATACCA

TGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCATGCAGCAAAT

GTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTCTTTCACCAAC

CTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTGGAGTCCACCCG

TTGTTTGCTGGCTGGGCTTTTCCAGTGTCAGAAAGAAGGACCCATCATCATCCACACTGATGA

AGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGAGGCAGAGAACCAG

AGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTGAAAAAGGTGAAGGA

CAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCTGGACAACGTGGCTGC

CGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTGCACGGATGATCGAACA

GAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGGGAAAGTCTTCAGATGGC

AGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGCCATGGACTCTGCCACTGC

CCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATGTGACCTTCATACCGCTCTT

AATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCTGTTGACCTGACCATGGAACT

TTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTATTACCACGGGAAGGAGCAGGT

GCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGTTCTTGAATGCCATGTCAGTTTA

TACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAAACTCAGGTGATGGAAGTTGGAAA

TGAAGAG**TAA**CTGATTTATAAAAGCAGGATGTGTTGATTTTAAAATAAAGTGCCTTTATACAATG

FIGURE 330

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGARSPLKPLPLEEQVE

WNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQLTKVGMQQMFALGERLRKNYVEDIPFL

SPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEGPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQ

PGISEDLKKVKDRMGIDSSDKVDFFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMA

VGPFLHILESNLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEWFVQ

LYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPEKYHALCSQTQVMEVGNEE

Important features:

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

FIGURE 331

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACATTTTTCCCTTTCCT
AACAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTTGCTGGAGAAGAAAGGGCTGAGGGCAG
AGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTGCCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGAT
GCAGCGGAGTGAGGTGATGGAAGTCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGA
CCTGGAGAGCCTGGGGGAGGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGATAAG
ACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGGTGAACTTCAACAGCCT
TTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAATAGAGATGCTTTGTAAAATAAAATTTTAA
AAAAAGCAAGTATTTTATAGCATAAAGGCTAGAGACCAAAATAGATAACAGGATTCCCTGAACATTCCTAAGAGG
GAGAAAGTATGTTAAAAATAGAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGG**ATG**GGGACC
CTGGGTCAGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTTGGTCTGACCACTCTGCCTTGTGTTTTGCAGAA
TCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCTCCTCACCGCCGCCCCTCTC
AGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGGACAGCCGCTCTGTGGTCCTGCTCTCAGTGGTC
TGGGTGCTGCTGGCCCCCCCAGCAGCCGGCATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACC
TTCAACCACTTGACCGTCCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACA
GGCAACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCGCCCCTCATC
GTGCAGCCCTGCAGCGAAGTGCTCACCCTCACCAACAATGTCAACAAGCTGCTCATCATTGACTACTCTGAGAAC
CGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCAAGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTG
GAGCCATCCCACAAGAAGGAGCACTACCTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGC
TCTGAGGGTGAGGATGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCC
AGCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGATTTTGTCTCCTCT
CTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACATCTTCTACATCTACGGCTTTGCTAGT
GGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGACCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTC
TTCTACACCTCACGCATCGTGCGGCTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGC
TGCACCCGGGCCGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCCAG
GCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAGCAGTATCACCACCCG
CCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTTGCAGATCAAGGAGCGCCTGCAGTCCTGC
TACCAGGGCGAGGGCAACCTGGAGCTCAACTGGCTGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCC
ATCGATGATAACTTCTGTGGACTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTG
TACACCACCAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTGTGGGG
ACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCCATTCACCTCCTCAGCAAA
GAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAGGCAACTTTATTTTCTTGGGGAACAAAGG**TGA**
AATGGGGAGGTAAGAAGGGGTTAATTTTGTGACTTAGCTTCTAGCTACTTCCTCCAGCCATCAGTCATTGGGTAT
GTAAGGAATGCAAGCGTATTTCAATATTTCCCAAACTTTAAGAAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

•

FIGURE 332

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWPR
ALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRVYK
LTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLACGSLYQG
VCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDYFPTL
SSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFLTVQPE
TPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKPGDSLAQ
AFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQIKERLQSCYQGEGNLELNWLL
GKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNGYSVVFVGT
KSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

Important features:

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 333

GCTGAGTCTGCTGCTCCTGCTGCTGCTGCTCCAGCCTGTAACCTGTGCCTACACCACGCCAGG

CCCCCCCAGAGCCCTCACCACGCTGGGCGCCCCCAGAGCCCACACC**ATG**CCGGGCACCTACGC

TCCCTCGACCACACTCAGTAGTCCCAGCACCCAGGGCCTGCAAGAGCAGGCACGGGCCCTGAT

GCGGGACTTCCCGCTCGTGGACGGCCACAACGACCTGCCCCTGGTCCTAAGGCAGGTTTACCA

GAAAGGGCTACAGGATGTTAACCTGCGCAATTTCAGCTACGGCCAGACCAGCCTGGACAGGCT

TAGAGATGGCCTCGTGGGCGCCCAGTTCTGGTCAGCCTATGTGCCATGCCAGACCCAGGACCG

GGATGCCCTGCGCCTCACCCTGGAGCAGATTGACCTCATACGCCGCATGTGTGCCTCCTATTC

TGAGCTGGAGCTTGTGACCTCGGCTAAAGCTCTGAACGACACTCAGAAATTGGCCTGCCTCAT

CGGTGTAGAGGGTGGCCACTCGCTGGACAATAGCCTCTCCATCTTACGTACCTTCTACATGCT

GGGAGTGCGCTACCTGACGCTCACCCACACCTGCAACACACCCTGGGCAGAGAGCTCCGCTAA

GGGCGTCCACTCCTTCTACAACAACATCAGCGGGCTGACTGACTTTGGTGAGAAGGTGGTGGC

AGAAATGAACCGCCTGGGCATGATGGTAGACTTATCCCATGTCTCAGATGCTGTGGCACGGCG

GGCCCTGGAAGTGTCACAGGCACCTGTGATCTTCTCCCACTCGGCTGCCCGGGGTGTGTGCAA

CAGTGCTCGGAATGTTCCTGATGACATCCTGCAGCTTCTGAAGAAGAACGGTGGCGTCGTGAT

GGTGTCTTTGTCCATGGGAGTAATACAGTGCAACCCATCAGCCAATGTGTCCACTGTGGCAGA

TCACTTCGACCACATCAAGGCTGTCATTGGATCCAAGTTCATCGGGATTGGTGGAGATTATGA

TGGGGCCGGCAAATTCCCTCAGGGGCTGGAAGACGTGTCCACATACCCGGTCCTGATAGAGGA

GTTGCTGAGTCGTGGCTGGAGTGAGGAAGAGCTTCAGGGTGTCCTTCGTGGAAACCTGCTGCG

GGTCTTCAGACAAGTGGAAAAGGTACAGGAAGAAACAAATGGCAAAGCCCCTTGGAGGACAA

GTTCCCGGATGAGCAGCTGAGCAGTTCCTGCCACTCCGACCTCTCACGTCTGCGTCAGAGACA

GAGTCTGACTTCAGGCCAGGAACTCACTGAGATTCCCATACACTGGACAGCCAAGTTACCAGC

CAAGTGGTCAGTCTCAGAGTCCTCCCCCCACATGGCCCCAGTCCTTGCAGTTGTGGCCACCTT

CCCAGTCCTTATTCTGTGGCTC**TGA**TGACCCAGTTAGTCCTGCCAGATGTCACTGTAGCAAGC

CACAGACACCCCACAAAGTTCCCCTGTTGTGCAGGCACAAATATTTCCTGAAATAAATGTTTT

GGACATAG

FIGURE 334

MPGTYAPSTTLSSPSTQGLQEQARALMRDFPLVDGHNDLPLVLRQVYQKGLQDVNLRNFSYGQ

TSLDRLRDGLVGAQFWSAYVPCQTQDRDALRLTLEQIDLIRRMCASYSELELVTSAKALNDTQ

KLACLIGVEGGHSLDNSLSILRTFYMLGVRYLTLTHTCNTPWAESSAKGVHSFYNNISGLTDF

GEKVVAEMNRLGMMVDLSHVSDAVARRALEVSQAPVIFSHSAARGVCNSARNVPDDILQLLKK

NGGVVMVSLSMGVIQCNPSANVSTVADHFDHIKAVIGSKFIGIGGDYDGAGKFPQGLEDVSTY

PVLIEELLSRGWSEEELQGVLRGNLLRVFRQVEKVQEENKWQSPLEDKFPDEQLSSSCHSDLS

RLRQRQSLTSGQELTEIPIHWTAKLPAKWSVSESSPHMAPVLAVVATFPVLILWL

Important features:

N-glycosylation sites.

amino acids 58-62, 123-127, 182-186, 273-277

N-myristoylation sites.

amino acids 72-78, 133-139, 234-240, 264-270, 334-340, 389-395

Renal dipeptidase active site.

amino acids 134-157

FIGURE 335

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCAG
AGCAGCCGGCTGCCGCCCCGGGAAG**ATG**GCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTGCT
GCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAAAGA
CCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAAAGAA
GACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTACTATCA
ACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCGGATCAA
AAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTGAGCAAGG
CCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTTCCATCATG
TGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGACAAAGAAGG
GAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAAATCCCAGACT
TGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAAACTGGAACTCTGCAATTTAA
TACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTCTGTTGGATATCG
CAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCATCATAGCAGCCGT
AGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTATGCTCAGAGGAAAGG
CTACTTTTCAAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAAAGCCACGACAATGAG
TGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGGCCGCGGCGGGCGGATC
ACGAGGTCAGGAGTTC**TAG**ACCAGTCTGGCCAATATGGTGAAACCCCATCTCTACTAAAATAC
AAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGCTTGGGAGACAGGAGAATC
ACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGCCACTGCAGTCCAGCCTGGG
TAACAGAGCAAGATTCCATCTCAAAAAATAAAATAAATAAATAAATACTGGTTTTTACC
TGTAGAATTCTTACAATAAATATAGCTTGATATTC

FIGURE 336

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLEW
KKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEEDTV
TLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQSTNSS
YTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYRRCPGKRMQVDDLNISGIIAAVVVVALVIS
VCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQEF

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

FIGURE 337

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCGG

CACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGCC**A**

TGATTTCCCTCCCGGGGCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGAGTG

CCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAGGCGG

TGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGCACGGGGAGGTGTCTTCATCCC

AGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATCAGGTGT

TGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCCATGCCCT

CCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTACAGCTGCT

CCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCTTAGAACTCA

ATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCATGTGGGGGCAA

ACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCAGTGGGATCGGC

AGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTGGGTCTTTAAGCC

TCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCACAATGAGGTGGGCA

CTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGCAGTGGTTGCTGGAG

CTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCCTCTTGTACCACCGCC

GGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCCATTGCTCCCCGGACCC

TGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCTTTCCTCTGTCACCTCCG

CACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCATTGACCCCCACGCCCAGTC

TCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCCACGACAGATGGGGCCCACCCTCAACCAA

TATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCGCATGGGTGCTGTGCCTGTGA

TGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTA**TGA**TGACCCCACCACTCATTGGCTAAAG

GATTTGGGGTCTCTCCTTCCTATAAGGGTCACCTCTAGCACAGAGGCCTGAGTCATGGGAAAG

AGTCACACTCCTGACCCTTAGTACTCTGCCCCCACCTCTCTTTACTGTGGGAAAACCATCTCA

GTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGAAGTGGATCTGGAATTGGGAGGAGC

CTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTGCTGAAATTAGCTACTCACCAAGAGT

GAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGCCCCCTTGATCTGTACCCCACCCCTAT

CTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGTATTGATATAACCTGTCAGGCTGGCTTG

GTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTCTTATTAAAACTAACATGAAATATGTGTT

GTTTTCATTTGCAAATTTAAATAAAGATACATAATGTTTGTATGAAAAA

FIGURE 338

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSSS
QPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPYSC
SVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSCQSPRSKPAVQYQWDR
QLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAAVVAG
AVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPWPKSSDTISKNGTLSSVTS
ARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSRMGAVPV
MVPAQSQAGSLV

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262, 262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 339

```
GCGAGAACCTTTGCACGCGCACAAACTACGGGGACGATTTCTGATTGATTTTTGGCGCTTTCGATCCACCCTCCT
CCCTTCTC**ATG**GGACTTTGGGGACAAAGCGTCCCGACCGCCTCGAGCGCTCGAGCAGGGCGCTATCCAGGAGCCA
GGACAGCGTCGGGAACCAGACCATGGCTCCTGGACCCCAAGATCCTTAAGTTCGTCGTCTTCATCGTCGCGGTTC
TGCTGCCGGTCCGGGTTGACTCTGCCACCATCCCCCGGCAGGACGAAGTTCCCCAGCAGACAGTGGCCCCACAGC
AACAGAGGCGCAGCCTCAAGGAGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAATATACTGGAGCCTGTAACC
CGTGCACAGAGGGTGTGGATTACACCATTGCTTCCAACAATTTGCCTTCTTGCCTGCTATGTACAGTTTGTAAAT
CAGGTCAAACAAATAAAAGTTCCTGTACCACGACCAGAGACACCGTGTGTCAGTGTGAAAAAGGAAGCTTCCAGG
ATAAAAACTCCCCTGAGATGTGCCGGACGTGTAGAACAGGGTGTCCCAGAGGGATGGTCAAGGTCAGTAATTGTA
CGCCCCGGAGTGACATCAAGTGCAAAAATGAATCAGCTGCCAGTTCCACTGGGAAAACCCCAGCAGCGGAGGAGA
CAGTGACCACCATCCTGGGGATGCTTGCCTCTCCCTATCACTACCTTATCATCATAGTGGTTTTAGTCATCATTT
TAGCTGTGGTTGTGGTTGGCTTTTCATGTCGGAAGAAATTCATTTCTTACCTCAAAGGCATCTGCTCAGGTGGTG
GAGGAGGTCCCGAACGTGTGCACAGAGTCCTTTTCCGGCGGCGTTCATGTCCTTCACGAGTTCCTGGGGCGGAGG
ACAATGCCCGCAACGAGACCCTGAGTAACAGATACTTGCAGCCCACCCAGGTCTCTGAGCAGGAAATCCAAGGTC
AGGAGCTGGCAGAGCTAACAGGTGTGACTGTAGAGTCGCCAGAGGAGCCACAGCGTCTGCTGGAACAGGCAGAAG
CTGAAGGGTGTCAGAGGAGGAGGCTGCTGGTTCCAGTGAATGACGCTGACTCCGCTGACATCAGCACCTTGCTGG
ATGCCTCGGCAACACTGGAAGAAGGACATGCAAAGGAAACAATTCAGGACCAACTGGTGGGCTCCGAAAAGCTCT
TTTATGAAGAAGATGAGGCAGGCTCTGCTACGTCCTGCCTG**TGA**AAGAATCTCTTCAGGAAACCAGAGCTTCCCT
CATTTACCTTTTCTCCTACAAAGGGAAGCAGCCTGGAAGAAACAGTCCAGTACTTGACCCATGCCCCAACAAACT
CTACTATCCAATATGGGGCAGCTTACCAATGGTCCTAGAACTTTGTTAACGCACTTGGAGTAATTTTTATGAAAT
ACTGCGTGTGATAAGCAAACGGGAGAAATTTATATCAGATTCTTGGCTGCATAGTTATACGATTGTGTATTAAGG
GTCGTTTTAGGCCACATGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGATAGGCTGAGGCAGGTGGATTGCTT
GAGCTCGGGAGTTTGAGACCAGCCTCATCAACACAGTGAAACTCCATCTCAATTTAAAAAGAAAAAAAGTGGTTT
TAGGATGTCATTCTTTGCAGTTCTTCATCATGAGACAAGTCTTTTTTTTCTGCTTCTTATATTGCAAGCTCCATCT
CTACTGGTGTGTGCATTTAATGACATCTAACTACAGATGCCGCACAGCCACAATGCTTTGCCTTATAGTTTTTTA
ACTTTAGAACGGGATTATCTTGTTATTACCTGTATTTTCAGTTTCGGATATTTTTGACTTAATGATGAGATTATC
AAGACGTAGCCCTATGCTAAGTCATGAGCATATGGACTTACGAGGGTTCGACTTAGAGTTTTGAGCTTTAAGATA
GGATTATTGGGGCTTACCCCCACCTTAATTAGAGAAACATTTATATTGCTTACTACTGTAGGCTGTACATCTCTT
TTCCGATTTTTGTATAATGATGTAAACATGGAAAAACTTTAGGAAATGCACTTATTAGGCTGTTTACATGGGTTG
CCTGGATACAAATCAGCAGTCAAAAATGACTAAAAATATAACTAGTGACGGAGGGAGAAATCCTCCCTCTGTGGG
AGGCACTTACTGCATTCCAGTCTTCCCTCCTGCGCCCTGAGACTGGACCAGGGTTTGATGGCTGGCAGCTTCTCA
AGGGGCAGCTTGTCTTACTTGTTAATTTTAGAGGTATATAGCCATATTTATTTATAAATAAATATTTATTTATTT
ATTTATAAGTAGATGTTTACATATGCCCAGGATTTTGAAGAGCCTGGTATCTTTGGGAAGCCATGTGTCTGGTTT
GTCGTGCTGGGACAGTCATGGGACTGCATCTTCCGACTTGTCCACAGCAGATGAGGACAGTGAGAATTAAGTTAG
ATCCGAGACTGCGAAGAGCTTCTCTTTCAAGCGCCATTACAGTTGAACGTTAGTGAATCTTGAGCCTCATTTGGG
CTCAGGCAGAGCAGGTGTTTATCTGCCCCGGCATCTGCCATGGCATCAAGAGGGAAGAGTGGACGGTGCTTGGG
AATGGTGTGAAATGGTTGCCGACTCAGGCATGGATGGGCCCCTCTCGCTTCTGGTGGTCTGTGAACTGAGTCCCT
GGGATGCCTTTTAGGGCAGAGATTCCTGAGCTGCGTTTTAGGGTACAGATTCCCTGTTTGAGGAGCTTGGCCCCT
CTGTAAGCATCTGACTCATCTCAGAGATATCAATTCTTAAACACTGTGACAACGGGATCTAAAATGGCTGACACA
TTTGTCCTTGTGTCACGTTCCATTATTTTATTTAAAAACCTCAGTAATCGTTTTAGCTTCTTTCCAGCAAACTCT
TCTCCACAGTAGCCCAGTCGTGGTAGGATAAATTACGGATATAGTCATTCTAGGGGTTTCAGTCTTTTCCATCTC
AAGGCATTGTGTGTTTTGTTCCGGGACTGGTTTGGCTGGGACAAAGTTAGAACTGCCTGAAGTTCGCACATTCAG
ATTGTTGTGTCCATGGAGTTTTAGGAGGGGATGGCCTTTCCGGTCTTCGCACTTCCATCCTCTCCCACTTCCATC
TGGCGTCCCACACCTTGTCCCCTGCACTTCTGGATGACACAGGGTGCTGCTGCCTCCTAGTCTTTGCCTTTGCTG
GGCCTTCTGTGCAGGAGACTTGGTCTCAAAGCTCAGAGAGAGCCAGTCCGGTCCCAGCTCCTTTGTCCCTTCCTC
AGAGGCCTTCCTTGAAGATGCATCTAGACTACCAGCCTTATCAGTGTTTAAGCTTATTCCTTTAACATAAGCTTC
CTGACAACATGAAATTGTTGGGGTTTTTTGGCGTTGGTTGATTTGTTTAGGTTTTGCTTTATACCCGGGCCAAAT
AGCACATAACACCTGGTTATATATGAAATACTCATATGTTTATGACCAAAATAAATATGAAACCTCATRTTAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 340

MGLWGQSVPTASSARAGRYPGARTASGTRPWLLDPKILKFVVFIVAVLLPVRVDSATIPRQDEVPQQTVAPQQQR
RSLKEEECPAGSHRSEYTGACNPCTEGVDYTIASNNLPSCLLCTVCKSGQTNKSSCTTTRDTVCQCEKGSFQDKN
SPEMCRTCRTGCPRGMVKVSNCTPRSDIKCKNESAASSTGKTPAAEETVTTILGMLASPYHYLIIIVVLVIILAV
VVVGFSCRKKFISYLKGICSGGGGGPERVHRVLFRRRSCPSRVPGAEDNARNETLSNRYLQPTQVSEQEIQGQEL
AELTGVTVESPEEPQRLLEQAEAEGCQRRRLLVPVNDADSADISTLLDASATLEEGHAKETIQDQLVGSEKLFYE
EDEAGSATSCL

Important features:
Transmembrane domains:
amino acids 35-52, 208-230

N-glycosylation sites.
amino acids 127-131, 182-186, 277-281

Glycosaminoglycan attachment site.
amino acids 245-249

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 260-264

N-myristoylation sites.
amino acids 21-27, 86-92, 102-108, 161-167, 242-248, 270-276, 297-303, 380-386

ATP/GTP-binding site motif A (P-loop).
amino acids 185-193

TNFR/NGFR cysteine-rich region.
amino acids 99-139

FIGURE 341

GCCTCTGAATTGTTGGGCAGTCTGGCAGTGGAGCTCTCCCCGGTCTGACAGCCACTCCAGAGG

CC**ATG**CTTCGTTTCTTGCCAGATTTGGCTTTCAGCTTCCTGTTAATTCTGGCTTTGGGCCAGG

CAGTCCAATTTCAAGAATATGTCTTTCTCCAATTTCTGGGCTTAGATAAGGCGCCTTCACCCC

AGAAGTTCCAACCTGTGCCTTATATCTTGAAGAAAATTTTCCAGGATCGCGAGGCAGCAGCGA

CCACTGGGGTCTCCCGAGACTTATGCTACGTAAAGGAGCTGGGCGTCCGCGGGAATGTACTTC

GCTTTCTCCCAGACCAAGGTTTCTTTCTTTACCCAAAGAAAATTTCCCAAGCTTCCTCCTGCC

TGCAGAAGCTCCTCTACTTTAACCTGTCTGCCATCAAAGAAAGGGAACAGTTGACATTGGCCC

AGCTGGGCCTGGACTTGGGGCCCAATTCTTACTATAACCTGGGACCAGAGCTGGAACTGGCTC

TGTTCCTGGTTCAGGAGCCTCATGTGTGGGGCCAGACCACCCCTAAGCCAGGTAAAATGTTTG

TGTTGCGGTCAGTCCCATGGCCACAAGGTGCTGTTCACTTCAACCTGCTGGATGTAGCTAAGG

ATTGGAATGACAACCCCCGGAAAAATTTCGGGTTATTCCTGGAGATACTGGTCAAAGAAGATA

GAGACTCAGGGGTGAATTTTCAGCCTGAAGACACCTGTGCCAGACTAAGATGCTCCCTTCATG

CTTCCCTGCTGGTGGTGACTCTCAACCCTGATCAGTGCCACCCTTCTCGGAAAAGGAGAGCAG

CCATCCCTGTCCCCAAGCTTTCTTGTAAGAACCTCTGCCACCGTCACCAGCTATTCATTAACT

TCCGGGACCTGGGTTGGCACAAGTGGATCATTGCCCCCAAGGGGTTCATGGCAAATTACTGCC

ATGGAGAGTGTCCCTTCTCACTGACCATCTCTCTCAACAGCTCCAATTATGCTTTCATGCAAG

CCCTGATGCATGCCGTTGACCCAGAGATCCCCCAGGCTGTGTGTATCCCCACCAAGCTGTCTC

CCATTTCCATGCTCTACCAGGACAATAATGACAATGTCATTCTACGACATTATGAAGACATGG

TAGTCGATGAATGTGGGTGTGGG**TAG**GATGTCAGAAATGGGAATAGAAGGAGTGTTCTTAGGG

TAAATCTTTTAATAAAACTACCTATCTGGTTTATGACCACTTAGATCGAAATGTC

FIGURE 342

MLRFLPDLAFSFLLILALGQAVQFQEYVFLQFLGLDKAPSPQKFQPVPYILKKIFQDREAAAT
TGVSRDLCYVKELGVRGNVLRFLPDQGFFLYPKKISQASSCLQKLLYFNLSAIKEREQLTLAQ
LGLDLGPNSYYNLGPELELALFLVQEPHVWGQTTPKPGKMFVLRSVPWPQGAVHFNLLDVAKD
WNDNPRKNFGLFLEILVKEDRDSGVNFQPEDTCARLRCSLHASLLVVTLNPDQCHPSRKRRAA
IPVPKLSCKNLCHRHQLFINFRDLGWHKWIIAPKGFMANYCHGECPFSLTISLNSSNYAFMQA
LMHAVDPEIPQAVCIPTKLSPISMLYQDNNDNVILRHYEDMVVDECGCG

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 112-116, 306-310

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 96-100

N-myristoylation site.

amino acids 77-83

TGF-beta family proteins.

amino acids 264-299, 327-341, 345-364

FIGURE 343

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCCATTCCTTTTCATTGACAAACTGACTTTTTTTATTTCT
TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT
CTGTGTTTGGGGTTTCTTCTTCCTCCCCTGACATTGGCATTGCTTAGTGGTTGTGTGGGGAGGGAGACCACGTGG
GCTCAGTGCTTGCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC
ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCGTGCTCTGTCTTTACTTCAAAATACACAAC
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG
AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC
AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG
CTCCCCATCAGTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAGAGCTTC
CACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCTCACTGTGGGGAATGCTGATAAACCAGTCA
CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAACTGACGTTTCCCTGGAGGTGTCCAGAAA
GCTGATGTAACACAGAGCCTATAAAAGCTGTCGGTCCTTAAGGCTGCCCAGCGCCTTGCCAAA**ATG**GAGCTTGTA
AGAAGGCTCATGCCATTGACCCTCTTAATTCTCTCCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT
GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC
AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAATTATCTTTTCCTATGTCCAGCTT
GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG
CAAGTCTGCAGTAAAAACGACTATGTTCCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCCTAACATCTCTATTCCAAACTGT
GGCGGTTACCTGGATACCTTGGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTTCAAAGAGATTTTCCTAGAAATAGAC
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT
GGCCGTGTGACTCCCACCTTCGAATCGTCATCAAACTCTCTGACTGTCGTGTTGTCTACAGATTATGCCAATTCT
TACCGGGGATTTTCTGCTTCCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA
GACCCAACTTGCAGACCAAAATTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC
CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA
GAAGATGATGTAATACAAAGTCAAAATGCACTGGGCAAATATAACACCAGCATGGCTCTTTTTGAATCCAATTCA
TTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACTCTTTTTGTTCAAGTTAGTCTGCAC
ACCTCAGATCCAAATTTGGTGGTGTTTCTTGATACCTGTAGAGCCTCTCCCACCTCTGACTTTGCATCTCCAACC
TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAAACTTGTAAGGTGTATCCCTTATTTGGACACTATGGGAGA
TTCCAGTTTAATGCCTTTAAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTTGATATGTGAT
AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG
AAAACAGATTCCATCATAGGACCCATTCGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTCAGCAT
GAAACACATGCGGAAGAAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCCTTCATGGTTCTAGCTCTG
AATGTGGTGACTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG
CAGAACTAT**TAA**CTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT
GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

FIGURE 344

MELVRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTIE
RPENKSIRIIFSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSSSTLTFQ
IVTDSARIQRTVFVFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQVEKD
YKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLSTDYA
NSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRPKLSNV
VEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNSTVEIIYI
TEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPNLVVFLDT
CRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYLQCKVLICD
SSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDRSASGNSGFQHETHAEETPNQPFN
SVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

Important features:
Signal sequence:
amino acids 1-24

Transmembrane domain:
amino acids 571-586

N-glycosylation site.
amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,
394-398, 419-423

Casein kinase II phosphorylation site.
amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,
408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.
amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.
amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 7-18

FIGURE 345

```
TGGGGGCCCCCCAGGCTCGCGCGTGGAGCGAAGCAGCATGGGCAGTCGGTGCGCGCTGGCCCTGGCGGTGCTCTC
GGCCTTGCTGTGTCAGGTCTGGAGCTCTGGGGTGTTCGAACTGAAGCTGCAGGAGTTCGTCAACAAGAAGGGGCT
GCTGGGGAACCGCAATTGCTGCCGCGGGGGCGCGGGGCCACCGCCGTGCGCCTGCCGGACCTTCTTCCGCGTGTG
CCTCAAGCACTACCAGGCCAGCGTGTCCCCCGAGCCGCCCTGCACCTACGGCAGCGCCGTCACCCCCGTGCTGGG
CGTCGACTCCTTCAGTCTGCCCGACGGCGGGGGCGCCGACTCCGCGTTCAGCAACCCCATCCGCTTCCCCTTCGG
CTTCACCTGGCCGGGCACCTTCTCTCTGATTATTGAAGCTCTCCACACAGATTCTCCTGATGACCTCGCAACAGA
AAACCCAGAAAGACTCATCAGCCGCCTGGCCACCCAGAGGCACCTGACGGTGGGCGAGGAGTGGTCCCAGGACCT
GCACAGCAGCGGCCGCACGGACCTCAAGTACTCCTACCGCTTCGTGTGTGACGAACACTACTACGGAGAGGGCTG
CTCCGTTTTCTGCCGTCCCCGGGACGATGCCTTCGGCCACTTCACCTGTGGGGAGCGTGGGGAGAAAGTGTGCAA
CCCTGGCTGGAAAGGGCCCTACTGCACAGAGCCGATCTGCCTGCCTGGATGTGATGAGCAGCATGGATTTTGTGA
CAAACCAGGGGAATGCAAGTGCAGAGTGGGCTGGCAGGGCCGGTACTGTGACGAGTGTATCCGCTATCCAGGCTG
TCTCCATGGCACCTGCCAGCAGCCCTGGCAGTGCAACTGCCAGGAAGGCTGGGGGGGCCTTTTCTGCAACCAGGA
CCTGAACTACTGCACACACCATAAGCCCTGCAAGAATGGAGCCACCTGCACCAACACGGGCCAGGGGAGCTACAC
TTGCTCTTGCCGGCCTGGGTACACAGGTGCCACCTGCGAGCTGGGGATTGACGAGTGTGACCCCAGCCCTTGTAA
GAACGGAGGGAGCTGCACGGATCTCGAGAACAGCTACTCCTGTACCTGCCCACCCGGCTTCTACGGCAAAATCTG
TGAATTGAGTGCCATGACCTGTGCGGACGGCCCTTGCTTTAACGGGGGTCGGTGCTCAGACAGCCCCGATGGAGG
GTACAGCTGCCGCTGCCCCGTGGGCTACTCCGGCTTCAACTGTGAGAAGAAAATTGACTACTGCAGCTCTTCACC
CTGTTCTAATGGTGCCAAGTGTGTGGACCTCGGTGATGCCTACCTGTGCCGCTGCCAGGCCGGCTTCTCGGGGAG
GCACTGTGACGACAACGTGGACGACTGCGCCTCCTCCCCGTGCGCCAACGGGGGCACCTGCCGGGATGGCGTGAA
CGACTTCTCCTGCACCTGCCCGCCTGGCTACACGGGCAGGAACTGCAGTGCCCCCGTCAGCAGGTGCGAGCACGC
ACCCTGCCACAATGGGGCCACCTGCCACGAGAGGGGCCACCGCTATGTGTGCGAGTGTGCCCGAGGCTACGGGGG
TCCCAACTGCCAGTTCCTGCTCCCCGAGCTGCCCCCGGGCCCAGCGGTGGTGGACCTCACTGAGAAGCTAGAGGG
CCAGGGCGGGCCATTCCCCTGGGTGGCCGTGTGCGCCGGGGTCATCCTTGTCCTCATGCTGCTGCTGGGCTGTGC
CGCTGTGGTGGTCTGCGTCCGGCTGAGGCTGCAGAAGCACCGGCCCCCAGCCGACCCCTGCCGGGGGGGAGACGGA
GACCATGAACAACCTGGCCAACTGCCAGCGTGAGAAGGACATCTCAGTCAGCATCATCGGGGCCACGCAGATCAA
GAACACCAACAAGAAGGCGGACTTCCACGGGGACCACAGCGCCGACAAGAATGGCTTCAAGGCCCGCTACCCAGC
GGTGGACTATAACCTCGTGCAGGACCTCAAGGGTGACGACACCGCCGTCAGGGACGCGCACAGCAAGCGTGACAC
CAAGTGCCAGCCCCAGGGCTCCTCAGGGGAGGAGAAGGGGACCCCGACCACACTCAGGGGTGGAGAAGCATCTGA
AAGAAAAAGGCCGACTCGGGCTGTTCAACTTCAAAAGACACCAAGTACCAGTCGGTGTACGTCATATCCGAGGA
GAAGGATGAGTGCGTCATAGCAACTGAGGTGTAAAATGGAAGTGAGATGGCAAGACTCCCGTTTCTCTTAAAATA
AGTAAAATTCCAAGGATATATGCCCCAACGAATGCTGCTGAAGAGGAGGGAGGCCTCGTGGACTGCTGCTGAGAA
ACCGAGTTCAGACCGAGCAGGTTCTCCTCCTGAGGTCCTCGACGCCTGCCGACAGCCTGTCGCGGCCCGGCCGCC
TGCGGCACTGCCTTCCGTGACGTCGCCGTTGCACTATGGACAGTTGCTCTTAAGAGAATATATATTTAAATGGGT
GAACTGAATTACGCATAAGAAGCATGCACTGCCTGAGTGTATATTTTGGATTCTTATGAGCCAGTCTTTTCTTGA
ATTAGAAACACAAACACTGCCTTTATTGTCCTTTTTGATACGAAGATGTGCTTTTTCTAGATGGAAAAGATGTGT
GTTATTTTTTGGATTTGTAAAAATATTTTTCATGATATCTGTAAAGCTTGAGTATTTTGTGATGTTCGTTTTTTA
TAATTTAAATTTTGGTAAATATGTACAAAGGCACTTCGGGTCTATGTGACTATATTTTTTTGTATATAAATGTAT
TTATGGAATATTGTGCAAATGTTATTTGAGTTTTTTACTGTTTTGTTAATGAAGAAATTCCTTTTTAAAATATTT
TTCCAAAATAAATTTTATGAATGACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAA
```

FIGURE 346

MGSRCALALAVLSALLCQVWSSGVFELKLQEFVNKKGLLGNRNCCRGGAGPPPCACRTFFRVC
LKHYQASVSPEPPCTYGSAVTPVLGVDSFSLPDGGGADSAFSNPIRFPFGFTWPGTFSLIIEA
LHTDSPDDLATENPERLISRLATQRHLTVGEEWSQDLHSSGRTDLKYSYRFVCDEHYYGEGCS
VFCRPRDDAFGHFTCGERGEKVCNPGWKGPYCTEPICLPGCDEQHGFCDKPGECKCRVGWQGR
YCDECIRYPGCLHGTCQQPWQCNCQEGWGGLFCNQDLNYCTHHKPCKNGATCTNTGQGSYTCS
CRPGYTGATCELGIDECDPSPCKNGGSCTDLENSYSCTCPPGFYGKICELSAMTCADGPCFNG
GRCSDSPDGGYSCRCPVGYSGFNCEKKIDYCSSSPCSNGAKCVDLGDAYLCRCQAGFSGRHCD
DNVDDCASSPCANGGTCRDGVNDFSCTCPPGYTGRNCSAPVSRCEHAPCHNGATCHERGHRYV
CECARGYGGPNCQFLLPELPPGPAVVDLTEKLEGQGGPFPWVAVCAGVILVLMLLLGCAAVVV
CVRLRLQKHRPPADPCRGETETMNNLANCQREKDISVSIIGATQIKNTNKKADFHGDHSADKN
GFKARYPAVDYNLVQDLKGDDTAVRDAHSKRDTKCQPQGSSGEEKGTPTTLRGGEASERKRPD
SGCSTSKDTKYQSVYVISEEKDECVIATEV

Important features:
Signal sequence:
Amino acids 1-21

Transmembrane domain:
Amino acids 546-566

N-glycosylation site:
Amino acids 477-481

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 660-664

Tyrosine kinase phosphorylation sites:
Amino acids 176-185;252-261

N-myristoylation sites:
Amino acids 2-8;37-43;40-46;98-104;99-105;262-268;281-287;
282-288;301-307;310-316;328-334;340-344;378-384;387-393;512-518;
676-682;683-689;695-701

Aspartic acid and asparagine hydroxylation sites:
Amino acids 343-355;420-432;458-470

Prokaryotic membrane lipoprotein lipid attachment site:
Amino acids 552-563

EGF-like domain cysteine pattern signature:
Amino acids 243-255;274-286;314-326;352-364;391-403;429-441;
467-479;505-517

FIGURE 347

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAACA
TCGTAGTCCACCCCCTCCCCATCCCCAGCCCCCGGGGATTCAGGCTCGCCAGCGCCCAGCCAG
GGAGCCGGCCGGGAAGCGCG**ATG**GGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGCTGT
TCGCCTGCTGCTGGGCGCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCTGGACAT
CTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCACGAGG
ACTCATCCCTGCAATGGTCTAACCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGAGAGCCC
TTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGCATCAGCA
ATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGTGCGAACTG
CCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAAGCCCATCATCACTGGTTATAAATCTT
CATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAGCCTGCAGCCC
GGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCATACAGGAAGATC
CCAATGGTAAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCCGGGAGGATGATG
GGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGACAGATCCACCTCTC
AACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCCTCCCCATCCTCGTG
AGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCCCCCAGCAGTACCTAT
GGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGTGCCCTGATCTTCCCTT
TCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAGCAACATGGGCAGCTACA
AGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCTCCTCCTCCAGCACCTACC
ACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTGCTCATCATGCTCATCTTCC
TTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACATGAGGCAAAAGGCTCCGACG
ATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCGGGCAGTCAGGAGGGGACGACA
AGAAGGAATATTTCATC**TAG**AGGCGCCTGCCCACTTCCTGCGCCCCCCAGGGGCCCTGTGGGG
ACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAACCGCAGGGCCGCCCCTCCCGCTT
GCTCCCCAGCCCACCCACCCCCCTGTACAGAATGTCTGCTTTGGGTGCGGTTTTGTACTCGGT
TTGGAATGGGGAGGGAGGAGGGCGGGGGGAGGGGAGGGTTGCCCTCAGCCCTTTCCGTGGCTT
CTCTGCATTTGGGTTATTATTATTTTTGTAACAATCCCAAATCAAATCTGTCTCCAGGCTGGA
GAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACAAACAAAAAACA

FIGURE 348

MGAPAASLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQWS
NPAQQTLYFGEKRALRDNRIQLVTSTPHELSISISNVALADEGEYTCSIFTMPVRTAKSLVTV
LGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGKTFT
VSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQKLLL
HCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKAYYTLN
VNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLTHEAKGSDDAPDADT
AIINAEGGQSGGDDKKEYFI

Important features:

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304, 306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 349

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAGGACAGCAGCAAAG

AGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTACCATACGCCCTCAGGACGTTCCCTCTA

GCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGTCATTTTGATTTTGCTGTTTATTTTTTTTTTCTTTTTCTT

TTTCCCACCACATTGTATTTTATTTCCGTACTTCAGAA**ATG**GGCCTACAGACCACAAAGTGGCCCAGCCATGGGG

CTTTTTTCCTGAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCCCTA

GTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTGCCTCTTGGGATCCCGG

AGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGCTGGATTTCCTGCAGAACTGCACAATGTAC

AGTCGGTGCACACGGTCTACCTGTATGGCAACCAACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAG

TTCTCCATTTGCAGGAAACAATATTCAGACCATTTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGC

TGCACCTGGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCCTCAAAT

TGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGACTTGCAAGAGCTGAGAGTGG

ATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAATCTCACGAGCTTGGAGCGTCTTATTGTGGACG

GGAACCTCCTGACCAACAAGGGTATCGCCGAGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTG

TACGTAATTCGCTGTCCCACCCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACC

AGATAAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCAACAACCAAC

TGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTCACTGCTCGGAATAACCCTTGGT

TTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAAATATATCCCTTCATCTCTCAACGTGCGGGGTTTCA

TGTGCCAAGGTCCTGAACAAGTCCGGGGGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCCTGTCCCACCA

CGACCCCCGGCCTGCCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTA

TTCCAAACCCTAGCAGAAGCTACACGCCTCCAACTCCTACCACATCGAAACTTCCCACGATTCCTGACTGGGATG

GCAGAGAAAGAGTGACCCCACCTATTTCTGAACGGATCCAGCTCTCTATCCATTTTGTGAATGATACTTCCATTC

AAGTCAGCTGGCTCTCTCTCTTCACCGTGATGGCATACAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAG

GGGGCATCGTTCAGGAGCGCATAGTCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCA

CCTATCGGATTTGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGCCA

CCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGACGTCCCACAGCATGG

GCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTTGTGCTGGTGGTCTTGCTCAGCGTCTTTT

GCTGGCATATGCACAAAAAGGGGCGCTACACCTCCCAGAAGTGGAAATACAACCGGGGCCGGCGGAAAGATGATT

ATTGCGAGGCAGGCACCAAGAAGGACAACTCCATCCTGGAGATGACAGAAACCAGTTTTCAGATCGTCTCCTTAA

ATAACGATCAACTCCTTAAAGGAGATTTCAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTACACAG

ACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGCACTGCCATACG**TGA**CAGC

CAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAACACACTCGTGTGTGCACATAAAGACACGCAG

ATTACATTTGATAAATGTTACACAGATGCATTTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAA

TGGGATTTAAAAAAAGTGCTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAA

TCTT

FIGURE 350

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIPE
GVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTISR
AALAQLLKLEELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPVDLQELRVDE
NRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPDLPGT
HLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNNPWFCD
CSIKWVTEWLKYIPSSLNVRGFMCQGPEQVRGMAVRELNMNLLSCPTTTPGLPLFTPAPSTAS
PTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVNDTSIQVSW
LSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPLDAFNYRAV
EDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVLLSVFCWHMH
KKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKGDFRLQPIYTP
NGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Important features:

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300, 522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 351

```
AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGCA
TGAAAGGGCTTCGCCGCCGGGAGTAAAAGAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC
GAGGGCGGGCGTGCACCCTCGGCTGGAAGTTTGTGCCGGGCCCCGAGCGCGCGCCGGCTGGGAGCTTCGGGTAGA
GACCTAGGCCGCTGGACCGCGATGAGCGCGCCGAGCCTCCGTGCGCGCGCCGCGGGGTTGGGGCTGCTGCTGTGC
GCGGTGCTGGGGCGCGCTGGCCGGTCCGACAGCGGCGGTCGCGGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC
GAGCGCCCATGCCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT
CCCGAGCCACTCCCGTCCTGGGTCGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC
ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA
GTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG
TTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATTTCCAGCCCTACAG
CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA
CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACTGCCCCAACTGCAA
CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT
CTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTG
CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACTTCAT
CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA
ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG
AACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC
AATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAGGCGACTGATACTC
CAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT
GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATTGCATTTAAATACA
TCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCCACAGTGGGTGGCGGAAAACAACTTTCAGAGCTTTGTA
AATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGATGGCTTTGTGTGT
GATGATTTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC
ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAATGAACTACTGCATGAT
GCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG
CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTCATCCTACTCTGTC
AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC
ATGGCACGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGATGGGGGCACAGAC
TTCCCAGCTGCACGGGAGAGACGCATGCCATGTGATGCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA
GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGCAACTCTGACTGTC
CTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGTCCTACAGTGC
ATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGTAACCGAGAGGCAC
TTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG
ATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAACCTGCGACTCCCCT
CAGATGACAGCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGCTGT
GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACC
AACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTTAGCTGACAGGCAG
GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGGATTTTTCTTACCA
CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTC
CTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGATCCTTTTGAAACA
TATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAG
GAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT
GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC
TCTTTAGATTTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA
GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCCAAGAGCCTTTTAT
TTGAAAGCTCATTCTTCCCCAGACTTGGACTCTGGGTCAGAGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAA
GAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC
ACATAGACTGAATGAGACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACTTTTATTTAAAAGAGAGAAT
CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAAATGCTTTATTTATACAGATGAACCAAAATTAC
AAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTATTTTTTAACTTTG
TTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTTATAATGCCAGA
TTTCTTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAAATAGAAGTT
ACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 352

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDCSRKRLARLPEPLPSW
VARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANITLLSLAGNRIVEILPEHLKEFQSLETL
DLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRN
KIKNVDGLTFQGLGALKSLKMQRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAIN
RISPDAWEFCQKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEISWTI
EDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMKKLQQLHLNTSSLLCDC
QLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDFPKPQITVQPETQSAIKGSNLSFICSAASS
SDSPMTFAWKKDNELLHDAEMENYAHLRAQGGEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVN
MLPSFTKTPMDLTIRAGAMARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYS
CTAQNSAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHFFAAGNQ
LLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWATVGVVIIAVVCCVVGTSLV
WVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLADRQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGT
CHIDNSSEADVEAATDLFLCPFLGSTGPMYLKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSH
PSEESCERSFSNISWPSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHL
DAYSSFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

Important features:
Signal sequence:
amino acids 1-27

Transmembrane domain:
amino acids 808-828

N-glycosylation site.
amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.
amino acids 886-890

Casein kinase II phosphorylation site.
amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,
1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.
amino acids 667-675

N-myristoylation site.
amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.
amino acids 58-80, 65-87

FIGURE 353

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCCAAACCCTTTTCTTCTCCTTTCCTGGCTTCGGACATTGG

AGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTGTTACTTTGTGATGAGATCGGGGATGA

ATTGCTCGCTTTAAAA**ATG**CTGCTTTGGATTCTGTTGCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTAC

AGGGGACGTTTGCAAAGAGAAGATCTGTTCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAA

GGGCTTCACAAGTCTGCAGCGTTTCACTGCCCCGACTTCCCAGTTTTACCATTTATTTCTGCATGGCAATTCCCT

CACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGGAAAACAATGGCTTGCA

TGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGGCTGCACATCAACAACAACAAGATCAAGTC

TTTTCGAAAGCAGACTTTTCTGGGGCTGGACGATCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATAT

AGACCCGGGGGCCTTCCAGGACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCTACC

TGCCAACGTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCCCTATGA

GGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTTGGGACTGCACCTGTGATCT

GCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCCCTGATCGGCCGAGTGGTCTGCGAAGCCCCCAC

CAGACTGCAGGGTAAAGACCTCAATGAAACCACCGAACAGGACTTGTGTCCTTTGAAAAACCGAGTGGATTCTAG

TCTCCCGGCGCCCCCTGCCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCA

AGAGGATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGATCAAAATCAG

ACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACAGTTTACCCTGCCCTGGGGGCTG

CAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAACTGCAACAACAGGAACGTGAGCAGCTTGGCTGATTT

GAAGCCCAAGCTCTCTAACGTGCAGGAGCTTTTCCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTT

TGTGGATTACAAGAACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAA

GAACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGGAGAAATTCGCGGG

GCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCATCCTCCCGGGCACTTTCAATGCCAT

GCCCAAACTGAGGATCCTCATTCTCAACAACAACCTGCTGAGGTCCCTGCCTGTGGACGTGTTCGCTGGGGTCTC

GCTCTCTAAACTCAGCCTGCACAACAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTC

CATCATCCAGATAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCAGA

ACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACTTCTTTAGAAAGGATTTCAT

GCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCTCGCCCACGTTAACTTCGCACAGTAAAAA

CAGCACTGGGTTGGCGGAGACCGGGACGCACTCCAACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTT

GGTCCCGGGACTGCTGCTGGTGTTTGTCACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTTTATCCTGAGGAA

CCGAAAGCGGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGTGACTC

TTCCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCCACAGAGTGTATGACTGTGGCTCTCACTCGCTCTC

AGAC**TAA**GACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCGATACATCCTTCCCCACCGCAGGCACCCCGGG

GGCTGGAGGGGCGTGTACCCAAATCCCCGCGCCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTC

GCACAACCGAAAGGGCCTGACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGA

GCGCAGCCAGCTCGCTCTTTGCTGAGAGCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAGAACTGACA

GTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATACATATATACATATATCCACATC

TATATAGAGAGATAGATATCTATTTTTCCCCTGTGGATTAGCCCCGTGATGGCTCCCTGTTGGCTACGCAGGGAT

GGGCAGTTGCACGAAGGCATGAATGTATTGTAAATAAGTAACTTTGACTTCTGAC

FIGURE 354

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYHL

FLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQTF

LGLDDLEYLQADFNLLRDIDPGAFQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRGNRL

KTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQGKDL

NETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTKIPGNW

QIKIRPTAAIATGSSRNKPLANSLPCPGGCSCDHIPGSGLKMNCNNRNVSSLADLKPKLSNVQ

ELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNYLDTLSRE

KFAGLQNLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSLSKLSLHNN

YFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKCETPVNFFRK

DFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVLVPGLLLVFVT

SAFTVVGMLVFILRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADGAHRVYDCGSHS

LSD

Important features:

Signal sequence:
amino acids 1-15

Transmembrane domain:
amino acids 618-638

N-glycosylation site.
amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 122-126, 646-650

Casein kinase II phosphorylation site.
amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577, 608-612, 657-661, 666-670, 693-697

N-myristoylation site.
amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349, 354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 337-348

FIGURE 355

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCCAGAATAACTCAGGGCTGCACCGGGCCTG

GCAGCGCTCCGCACACATTTCCTGTCGCGGCCTAAGGGAAACTGTTGGCCGCTGGGCCCGCGGGGGGATTCTTGG

CAGTTGGGGGGTCCGTCGGGAGCGAGGGCGGAGGGGAAGGGAGGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAG

GGCGGTGACCGCGCTCCAGACACAGCTCTGCGTCCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGC

GGGGCCTCAGAGA**ATG**AGGCCGGCGTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCGGGCCGGGCGGCGG

CGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGCACCACGCTACCATGAA

GCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTCAGCACCGTGCGTGCGGGCGCCGAGCTGCG

CGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGGGCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTCGCACT

GGAGCGCAGGCGTTCCCACTGCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCCTCCGACCCCGG

CGGTCTCGAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGCGGTACT

CCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGCGCGCCAACGGCTACCTGTG

CAAGTACCAGTTTGAGGTCTTGTGTCCTGCGCCGCGCCCCGGGGCCGCCTCTAACTTGAGCTATCGCGCGCCCTT

CCAGCTGCACAGCGCCGCTCTGGACTTCAGTCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCC

GATCTCAGTTACTTGCATCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTG

CCCCGGGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGGCTTTGCCTG

CGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCAGTGGGGAAGGACAGCCGACCCT

TGGGGGGACCGGGGTGCCCACCAGGCGCCCGCCGGCCACTGCAACCAGCCCCGTGCCGCAGAGAACATGGCCAAT

CAGGGTCGACGAGAAGCTGGGAGAGACACCACTTGTCCCTGAACAAGACAATTCAGTAACATCTATTCCTGAGAT

TCCTCGATGGGGATCACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCAC

CCCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTTCGACTCCTCCTC

TGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTTGGTGATCTTGACCATGACAGTACTGGGCTTGT

CAAGCTCTGCTTTCACGAAAGCCCCTCTTCCCAGCCAAGGAAGGAGTCTATGGGCCCGCCGGGCCTGGAGAGTGA

TCCTGAGCCCGCTGCTTTGGGCTCCAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCT

GCGGGACAGAGCAGAGGGTGCCTTGCTGGCGGAGTCCCCTCTTGGCTCTAGTGATGCA**TAG**GGAAACAGGGGACA

TGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTACTTGTGTAACTGACAA

TTTCTGCAGAAATCCCCCTTCCTCTAAATTCCCTTTACTCCACTGAGGAGCTAAATCAGAACTGCACACTCCTTC

CCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGATGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATT

TTCTTATGTTTATTCGGAGAATTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATAT

AATTTACATTAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTCAGGCTAGGAGTATATTGG

TTCGAAATCCCAGGGAAAAAAATAAAAATAAAAAATTAAAGGATTGTTGAT

FIGURE 356

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALST
VRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLESD
TLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCPAPRPGAASN
LSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPGRYLR
AGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPPATATSPVP
QRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITPSGSVIS
KFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMTVLGLVKLCFHESPSSQPRKESMGPP
GLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Important features:

Signal sequence:
amino acids 1-16

Transmembrane domain:
amino acids 399-418

N-glycosylation site.
amino acids 189-193, 381-385

Glycosaminoglycan attachment site.
amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 98-102, 434-438

Casein kinase II phosphorylation site.
amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.
amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157, 185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469, 477-483

Aspartic acid and asparagine hydroxylation site.
amino acids 262-274

FIGURE 357

```
CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGAAGA
AGACTAAAA**ATG**GTGTTTCCAATGTGGACACTGAAGAGACAAATTCTTATCCTTTTTAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTCCTAAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTTGACAGAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCCTC
ACCATTAACCACATACCAGACATCTCCCCAGCGTCCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTTAGTGGACTCACTTATTTAAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACAACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACTGTTATTATCGAAATCCTTGTTATGTTTCATATTCAATAGAGAAA
GATGCCTTCCTAAACTTGACAAAGTTAAAAGTGCTCTCCCTGAAAGATAACAATGTCACAGCCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
CTCAACCAATTACAAATTCTTGACCTAAGTGGAAATTGCCCTCGTTGTTATAATGCCCCATTTCCTTGTGCGCCG
TGTAAAAATAATTCTCCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAAAGTTTTACGTCTA
CACAGTAACTCTCTTCAGCATGTGCCCCCAAGATGGTTTAAGAACATCAACAAACTCCAGGAACTGGATCTGTCC
CAAAACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG
TCTTTCAATTTTGAACTTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAATTGCTAACCTCAGCAGTGTTTAAACAATTTAAAAGA
CTGAAAGTCATAGATCTTTCAGTGAATAAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAT
GCCAGAACTTCTGTAGAAAGTTATGAACCCCAGGTCCTGGAACAATTACATTATTTCAGATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTCATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCCTCTGATTTTCAGCATCTTTCTTTCCTCAAATGCCTG
AATCTGTCAGGAAATCTCATTAGCCAAACTCTTAATGGCAGTGAATTCCAACCTTTAGCAGAGCTGAGATATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTT
CTGCAGAAACTGATGATGAACGACAATGACATCTCTTCCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCCTAAGTTTCTTGCCTTCTGGAGTTTTTGATGGT
ATGCCTCCAAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAAACTCCAGTGT
CTAAAGAACCTGGAAACTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACTGTTCC
AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTC
AACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGG
GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTTCTCATGGTGATGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAA
CATTTTAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCCAGAGCATA
CAGCTTAGCAAAAGACAGTGTTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTTAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG
TCCAAGTTCCTCCAGCTCCGGAAAAGGCTCTGTGGGAGTTCTGTCCTTGAGTGGCCAACAAACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTC**TAG**CCCTTCTTTGCAAAACACAACTGCCTAGTTTACCAAGGAGAGGCCTGGC
```

FIGURE 358

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGGI
PTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFSGL
TYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYRNPC
YVSYSIEKDAFLNLTKLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNLNQLQ
ILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWFKNINK
LQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSLKILRIR
GYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKISPSGDSSE
VGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQTLDLSKNS
IFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLHSTAFEELHK
LEVLDISSNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSSTSRTMESESLRTLEFRGNH
LDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNLSLAKNGLKSFSWK
KLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQDAFQLRYLDLSS
NKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWWVNHTEVTIPYLATDVTCVGPGA
HKGQSVISLDLYTCELDLTNLILFSLSISVSLFLMVMMTASHLYFWDVWYIYHFCKAKIKGYQ
RLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCLEERDWLPGQPVLENLSQ
SIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFLEKPFQKSKFLQLRKRLC
GSSVLEWPTNPQAHPYFWQCLKNALATDNHVAYSQVFKETV

Important features:

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 359

GACGGCTGGCCACC**ATG**CACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTAC

TGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGTGG

AGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGAGAT

GGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCACAACA

AGGAGCGCGGGCGCCGCGGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGTGCCGC

TGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCTGCAGCC

CAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATCGGCTGTG

GTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACTGGTGTGCA

ACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTCCGTGCTCCC

AATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGCCCGGAAGATG

CTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGAAGCATCAGACT

CTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCTTGGTAACAGAGG

TCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAAACCCAGGCCCCAACTTCCTTAG

CAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAACAACTGAGGTCCCTT

CCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTACCTTCCCCAAATCGA

CCCATGTTCCTATCCCAAAATCAGCAGACAAAGTGACAGACAAAACAAAAGTGCCCTCTAGGA

GCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAGGGAACTCCTACCCCATG

CCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTGAGGTCTTGGCCTCAGTTT

TTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGACCACACGGGGCACACCTCCT

CCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAATGCCACGGGTGGGCGTGCCC

TGGCTCTGCAGTCGTCCTTGCCAGGTGCAGAGGGCCCTGACAAGCCTAGCGTTGTGTCAGGGC

TGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGACTACTGCTCCTGCCTCCTCTGG

TGTTGGCTGGAATCTTC**TGA**ATGGGATACCACTCAAAGGGTGAAGAGGTCAGCTGTCCTCCTG

TCATCTTCCCCACCCTGTCCCCAGCCCCTAAACAAGATACTTCTTGGTTAAGGCCCTCCGGAA

GGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCCATCCTGGAGGCACAAGGCCTGGCTG

GCTGCGAGCTCAGGAGGCCGCCTGAGGACTGCACACCGGGCCCACACCTCTCCTGCCCCTCCC

TCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCTCACTGCCTACCTGGCCTGGGGCTGTCT

GCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTGTGTAGCTGGGGATGGGGATTCCTAGGGG

CAGATGAAGGACAAGCCCCACTGGAGTGGGGTTCTTTGAGTGGGGGAGGCAGGGACGAGGGAA

GGAAAGTAACTCCTGACTCTCCAATAAAAACCTGTCCAACCTGTGAAA

FIGURE 360

MHGSCSFLMLLLPLLLLLVATTGPVGALTDEEKRLMVELHNLYRAQVSPTASDMLHMRWDEEL

AAFAKAYARQCVWGHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQMC

GHYTQVVWAKTERIGCGSHFCEKLQGVEETNIELLVCNYEPPGNVKGKRPYQEGTPCSQCPSG

YHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRATEASDSRKMGTPSSLATGIPAFLVTEVSGSL

ATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKSTHVPI

PKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAELPPSSEVLASVFPAQD

KPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVVSGLNSGP

GHVWGPLLGLLLLPPLVLAGIF

Important features:
Signal sequence:
amino acids 1-22

N-glycosylation site.
amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.
amino acids 439-443

Casein kinase II phosphorylation site.
amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.
amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

Amidation site.
amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.
amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.
amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.
amino acids 160-172

FIGURE 361

GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGG

GGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT

GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACTGCG

AGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCT**ATG**AGGCCACTCCTCGTCCTGCTGCT

CCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGG

GCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGG

CCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGG

ACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGG

GCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCG

GGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACA

TTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCA

TGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTC

TTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG

GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGC

CAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCC

AGTCTTTGCT**TAG**TGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGA

GGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGG

TGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCA

GGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGG

CAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC

CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGAT

CACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 362

MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPG
EKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFD
RVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPA
SLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA

Important features:

Signal sequence.

amino acids 1-15

N-myristoylation sites.

amino acids 11-17, 68-74, 216-222

Cell attachment sequence.

amino acids 77-80

FIGURE 363

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCG**ATG**ATGTGGCGACCATCAGTTCTGCTGCT
TCTGTTGCTACTGAGGCACGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGCCA
GGGGAGGGTGCACCAGGCGGCCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAACTT
CCAGTACGACCATGAGGCTTTCCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCACCCC
AGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGACGGCGA
CGGCTGGGTGTCGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCACATACG
GGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGGGTTGGGA
GGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTCATGACGTGGAGGA
TGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGCCGACCAGGA
TGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGGAGTTCCCTCA
CATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAAGATGGCTATGT
CCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGAGGAGCCGGCGTG
GGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGGATGGGCACCTGGA
TGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCCCTGGTGGAAGCCAA
CCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAAAGCGGAAATCCTGGG
TAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGGACCTGACCCGGCACCA
CGATGAGCTG**TGA**GCACCGCGCACCTGCCACAGCCTCAGAGGCCCGCACAATGACCGGAGGAG
GGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGCAGGAGGCAGATGCAGTCCCAGGC
ATCCTCCTGCCCCTGGGCTCTCAGGGACCCCCTGGGTCGGCTTCTGTCCCTGTCACACCCCCA
ACCCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAGCAATACCTATTTCTGACTGAGTCTCC
CAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAAGAACCGCCCCAACCCCTCCAG
CTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCTGGCCCCAGCCTCTCCTGCCTG
GCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAAAAGCCAGCGCCGGGACCTTGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 364

MMWRPSVLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGRE

VAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDTYD

TDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATREELT

AFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQFRDF

RDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFVGSQAT

NYGEDLTRHHDEL

Important features:

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293, 291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

FIGURE 365

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCG**ATG**GGGACAAAG

GCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGGCA

TTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCTGTG

AAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCAAGGA

GACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGGTGACC

TTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATACACTTGT

ATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGTGCTTGTG

CCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGGCAGTGCTG

ACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGATGGGATAGTG

ATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCTGAATCCCACA

ACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACAGCTGTGAGGCA

CGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCTGTGGAGCGGAAT

GTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAATCTTGGTTTTTGGC

ATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGACTTCGAGTAAGAAG

GTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAGACCTCGTCATTCCTG

GTG**TGA**GCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACTCAGGTGCTACCGGACT

CTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTCTACACCCCACAGGGCCC

CCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCCATCCTCCTTCATGCCCTC

CCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACTTGTTTAAAGTGTTTATTCCCCATTTCT

TTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCCTTCTAAGTAGACAGCAAAAA

TGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGCTGGCAGGGATCTTTGAATAGG

TATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGACGACCAGGGCCAGCTGTTCTAGA

GCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGGTGATGACACTGGGGTCCTTCCAT

CTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTGCCACTGGGATCCCTCTGCCCTGTC

CTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGTGGAAAATGGGAGCTCTTGTTGTGGA

GAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAGGATTTAAAACCGCTGCTCTAAAGAAA

AGAAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTGTAATCCCAGAGGCTGAGGCAGGCGGAT

CACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCAACATGGAGAAACCCTACTGGAAATACAA

AGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTCCCAGCTGCTCAGGAGCCTGGCAACAAGAG

CAAAACTCCAGCTCAAAAAAAAAAAAAAAA

FIGURE 366

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEWK

FDQGDTTRLVCYNNKITASYEDRVTFLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKVKL

IVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNSSYV

LNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVAAVLVTLILLGI

LVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Important features:

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158, 193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

FIGURE 367

```
GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG
AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGAAC
GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA
TAAACCAGAGTTAGACCCGCGGGGGTTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCCCTCC
CCACCCCCAAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTTTTCTC
TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT
GGTGTGGTGGTGTTTTCCTTTCTTTTTGAATTTCCCACAAGAGGAGAGGAAATTAATAATACATCTGCAAAGAAA
TTTCAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCACAGTTGGA
TTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTTTTTTTAAAT
TTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCTGGATGTTGCT
GTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGAATGTTGAACAAGAT
GACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCTGCTTGTGGTGCT
GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA
CCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT
GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT
ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA
ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTT
GCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTTGCGCCGACTAGACTTAGG
GGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACTTGAGGTATTTGAACCTTGC
CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAGCTGGATCTTTCTGGGAATCA
TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACTGTGGATGATACAGTCCCAGAT
TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC
ATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAACTG
TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTTGTGCCCGGTG
TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT
GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTCGGGCCTCCACATC
CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCGGATAGCTGT
GCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA
TTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTC
AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCC
AGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTCGACAGAGAAAACCTT
CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTCATGAAGACTACCAAAATCATCAT
TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA
TCGGCAAAACCATCACGCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC
CATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT
CAACCACACAACAACAGTTAACACAATAAATTCAATACACAGTTCAGTGCATGAACCGTTATTGATCCGAATGAA
CTCTAAAGACAATGTACAAGAGACTCAAATCTAAAACATTTACAGAGTTACAAAAAACAAACAATCAAAAAAAAA
GACAGTTTATTAAAAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAAACAA
AAAAGAAAAGAAATTTATTTATTAAAAATTCTATTGTGATCTAAAGCAGACAAAAA
```

FIGURE 368

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVICVRKNLREVPDGIS
TNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLANLNTLELFDNRLTTIPNGAFVYLSKL
KELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLSYISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELD
LSGNHLSAIRPGSFQGLMHLQKLWMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHH
NPWNCNCDILWLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAELKC
RASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGNTTASATLNVTAATTTP
FSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQSTRSTEKTFTIPVTDINSGIPGIDEVMKT
TKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQNHHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNS
YKSPFNHTTTVNTINSIHSSVHEPLLIRMNSKDNVQETQI

Important features:

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446,
488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397,
422-428, 433-439, 531-537

FIGURE 369

```
CAAAACTTGCGTCGCGGAGAGCGCCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC
TGGGGGAGCGCGTTCGGCCTGTGGGGCGCCGCTCGGCGCCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC
CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCCTGGGTGGTCCCGTCCCCTATCCCTCCTTTATATA
GAAACCTTCCACACTGGGAAGGCAGCGGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG
GCGCACAGCATTCCGAGTTTACAGATTTTTACAGATACCAA**ATG**GAAGGCGAGGAGGCAGAACAGCCTGCCTGGT
TCCATCAGCCCTGGCGCCCAGGCGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCCAGAGCCGGGTGCTGC
TGCTCCTGCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAA
GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGCTGGTACTGAGCCCTGAGG
AGCCCGGGCCTGGCCCAGCCGCGGTCAGCTGCCCCCGAGACTGTGCCTGTTCCCAGGAGGGCGTCGTGGACTGTG
GCGGTATTGACCTGCGTGAGTTCCCGGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACAACCAGC
TGGAAAAGATCTACCCTGAGGAGCTCTCCCGGCTGCACCGGCTGGAGACACTGAACCTGCAAAACAACCGCCTGA
CTTCCCGAGGGCTCCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC
TGACCTTGGCACCCCGCTTCCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCT
ATGGGCTCACCTTTGGCCAGAAGCCAAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGC
TGCCGGACAACATGTTCAACGGCTCCAGCAACGTCGAGGTCCTCATCCTGTCCAGCAACTTCCTGCGCCACGTGC
CCAAGCACCTGCCGCCTGCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCCGGGGGCCT
TCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGGCCTGGACAACGAGA
CCTTCTGGAAGCTCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCCAGCTGGGCTGC
CGCGCAGCCTGGTGCTGCTGCACTTGGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCCATCC
GCAGCCTGGAGTACCTGCTGCTGCACAGCAACCAGCTGCGGGAGCAGGGCATCCACCCACTGGCCTTCCAGGGCC
TCAAGCGGTTGCACACGGTGCACCTGTACAACAACGCGCTGGAGCGCGTGCCCAGTGGCCTGCCTCGCCGCGTGC
GCACCCTCATGATCCTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCCTGGAGG
AGCTCAACCTCAGCTACAACCGCATCACCAGCCCACAGGTGCACCGCGACGCCTTCCGCAAGCTGCGCCTGCTGC
GCTCGCTGGACCTGTCGGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGCTGAAGG
TCAAGCGCAATGAGCTGGCTGCCTTGGCACGAGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACCTCA
CCAGCAACCGACTGCGCAGCCGAGCCCTGGGCCCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA
TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAACA
AGATTAGTGCGGTGCCCGCCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAACAAGC
TGGCTGTGGGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT
TAGAGTTTGGTGACATTTCCAAGGACCGTGGCCGCTTGGGGAAGGAAAAGGAGGAGGAGGAAGAGGAGGAGGAGG
AGGAAGAGGAAACAAGA**TAG**TGACAAGGTGATGCAGATGTGACCTAGGATGATGGACCGCCGGACTCTTTTCTGC
AGCACACGCCTGTGTGCTGTGAGCCCCCCACTCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCA
TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCCACGGCGTGTCCCACGGCCAGACACATGC
ACACACATCACACCCTCAAACACCCAGCTCAGCCACACACAACTACCCTCCAAACCACCACAGTCTCTGTCACAC
CCCCACTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGGTCTGCCCCTGCCCTGGCACACACAGGCACCCA
TTCCCTCCCCCTGCTGACATGTGTATGCGTATGCATACACACCACACACACACATGCACAAGTCATGTGCGAA
CAGCCCTCCAAAGCCTATGCCACAGACAGCTCTTGCCCCAGCCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCT
GTCCATCTGTCCGTCCGTTCCCTGGAGAAGACACAAGGGTATCCATGCTCTGTGGCCAGGTGCCTGCCACCCTCT
GGAACTCACAAAAGCTGGCTTTTATTCCTTTCCCATCCTATGGGACAGGAGCCTTCAGGACTGCTGGCCTGGCC
TGGCCCACCCTGCTCCTCCAGGTGCTGGGCAGTCACTCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA
CAGGCACTTTTCCAATGGGCAAGCCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGCTGGGGCCTTGGGG
CAGGAGTGAAGCAGAGGTGATGGGGCTGGGCTGAGCCAGGGAGGAAGGACCCAGCTGCACCTAGGAGACACCTTT
GTTCTTCAGGCCTGTGGGGGAAGTTCCGGGTGCCTTTATTTTTTATTCTTTTCTAAGGAAAAAAATGATAAAAAT
CTCAAAGCTGATTTTTCTTGTTATAGAAAAACTAATATAAAAGCATTATCCCTATCCCTGCAAAAAAAAAA
```

FIGURE 370

MEGEEAEQPAWFHQPWRPGASDSAPPAGTMAQSRVLLLLLLLLPPQLHLGPVLAVRAPGFGRSG

GHSLSPEENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLPEH

TNHLSLQNNQLEKIYPEELSRLHRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNKLTL

APRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNVEVLI

LSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDNETFWK

LSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSNQLREQG

IHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYFLEELNLS

YNRITSPQVHRDAFRKLRLLRSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALARGALAGMA

QLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYLQNNKISAVP

ANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKDRGRLGKEKEE

EEEEEEEEEETR

Important features:

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341, 477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493, 535-557

FIGURE 371

```
CACTTTCTCCCTCTCTTCCTTTACTTTCGAGAAACCGCGCTTCCGCTTCTGGTCGCAGAGACCTCGGAGACCGCG
CCGGGGAGACGGAGGTGCTGTGGGTGGGGGGGACCTGTGGCTGCTCGTACCGCCCCCCACCCTCCTCTTCTGCAC
TGCCGTCCTCCGGAAGACCTTTTCCCCTGCTCTGTTTCCTTCACCGAGTCTGTGCATCGCCCCGGACCTGGCCGG
GAGGAGGCTTGGCCGGCGGGAGATGCTCTAGGGGCGGCGCGGGAGGAGCGGCCGGCGGGACGGAGGGCCCGGCAG
GAAG**ATG**GGCTCCCGTGGACAGGGACTCTTGCTGGCGTACTGCCTGCTCCTTGCCTTTGCCTCTGGCCTGGTCCT
GAGTCGTGTGCCCCATGTCCAGGGGGAACAGCAGGAGTGGGAGGGGACTGAGGAGCTGCCGTCGCCTCCGGACCA
TGCCGAGAGGGCTGAAGAACAACATGAAAAATACAGGCCCAGTCAGGACCAGGGGCTCCCTGCTTCCCGGTGCTT
GCGCTGCTGTGACCCCGGTACCTCCATGTACCCGGCGACCGCCGTGCCCCAGATCAACATCACTATCTTGAAAGG
GGAGAAGGGTGACCGCGGAGATCGAGGCCTCCAAGGGAAATATGGCAAAACAGGCTCAGCAGGGGCCAGGGGCCA
CACTGGACCCAAAGGGCAGAAGGGCTCCATGGGGGCCCCTGGGGAGCGGTGCAAGAGCCACTACGCCGCCTTTTC
GGTGGGCCGGAAGAAGCCCATGCACAGCAACCACTACTACCAGACGGTGATCTTCGACACGGAGTTCGTGAACCT
CTACGACCACTTCAACATGTTCACCGGCCAAGTTCTACTGCTACGTGCCCGGCCTCTACTTCTTCAGCCTCAACGT
GCACACCTGGAACCAGAAGGAGACCTACCTGCACATCATGAAGAACGAGGAGGAGGAGGTGGTGATCTTGTTCGCGCA
GGTGGGCGACCGCAGCATCATGCAAAGCCAGAGCCTGATGCTGGAGCTGCGAGAGCAGGACCAGGTGTGGGTACG
CCTCTACAAGGGCGAACGTGAGAACGCCATCTTCAGCGAGGAGCTGGACACCTACATCACCTTCAGTGGCTACCT
GGTCAAGCACGCCACCGAGCCC**TAG**CTGGCCGGCCACCTCCTTTCCTCTCGCCACCTTCCACCCCTGCGCTGTGC
TGACCCCACCGCCTCTTCCCCGATCCCTGGACTCCGACTCCCTGGCTTTGGCATTCAGTGAGACGCCCTGCACAC
ACAGAAAGCCAAAGCGATCGGTGCTCCCAGATCCCGCAGCCTCTGGAGAGAGCTGACGGCAGATGAAATCACCAG
GGCGGGGCACCCGCGAGAACCCTCTGGGACCTTCCGCGGCCCTCTCTGCACACATCCTCAAGTGACCCCGCACGG
CGAGACGCGGGTGGCGGCAGGGCGTCCCAGGGTGCGGCACCGCGGCTCCAGTCCTTGGAAATAATTAGGCAAATT
CTAAAGGTCTCAAAAGGAGCAAAGTAAACCGTGGAGGACAAAGAAAAGGGTTGTTATTTTTGTCTTTCCAGCCAG
CCTGCTGGCTCCCAAGAGAGAGGCCTTTTCAGTTGAGACTCTGCTTAAGAGAAGATCCAAAGTTAAAGCTCTGGG
GTCAGGGGAGGGGCCGGGGGCAGGAAACTACCTCTGGCTTAATTCTTTTAAGCCACGTAGGAACTTTCTTGAGGG
ATAGGTGGACCCTGACATCCCTGTGGCCTTGCCCAAGGGCTCTGCTGGTCTTTCTGAGTCACAGCTGCGAGGTGA
TGGGGGCTGGGGCCCCAGGCGTCAGCCTCCCAGAGGGACAGCTGAGCCCCCTGCCTTGGCTCCAGGTTGGTAGAA
GCAGCCGAAGGGCTCCTGACAGTGGCCAGGGACCCCTGGGTCCCCCAGGCCTGCAGATGTTTCTATGAGGGGCAG
AGCTCCTTGGTACATCCATGTGTGGCTCTGCTCCACCCCTGTGCCACCCCAGAGCCCTGGGGGGTGGTCTCCATG
CCTGCCACCCTGGCATCGGCTTTCTGTGCCGCCTCCCACACAAATCAGCCCCAGAAGGCCCCGGGGCCTTGGCTT
CTGTTTTTTATAAAACACCTCAAGCAGCACTGCAGTCTCCCATCTCCTCGTGGGCTAAGCATCACCGCTTCCACG
TGTGTTGTGTTGGTTGGCAGCAAGGCTGATCCAGACCCCTTCTGCCCCCACTGCCCTCATCCAGGCCTCTGACCA
GTAGCCTGAGAGGGGCTTTTTCTAGGCTTCAGAGCAGGGGAGAGCTGGAAGGGGCTAGAAAGCTCCCGCTTGTCT
GTTTCTCAGGCTCCTGTGAGCCTCAGTCCTGAGACCAGAGTCAAGAGGAAGTACACGTCCCAATCACCCGTGTCA
GGATTCACTCTCAGGAGCTGGGTGGCAGGAGAGGCAATAGCCCCTGTGGCAATTGCAGGACCAGCTGGAGCAGGG
TTGCGGTGTCTCCACGGTGCTCTCGCCCTGCCCATGGCCACCCCAGACTCTGATCTCCAGGAACCCCATAGCCCC
TCTCCACCTCACCCCATGTTGATGCCCAGGGTCACTCTTGCTACCCGCTGGGCCCCCAAACCCCCGCTGCCTCTC
TTCCTTCCCCCCATCCCCCACCTGGTTTTGACTAATCCTGCTTCCCTCTCTGGGCCTGGCTGCCGGGATCTGGGG
TCCCTAAGTCCCTCTCTTTAAAGAACTTCTGCGGGTCAGACTCTGAAGCCGAGTTGCTGTGGGCGTGCCCGGAAG
CAGAGCGCCACACTCGCTGCTTAAGCTCCCCCAGCTCTTTCCAGAAAACATTAAACTCAGAATTGTGTTTTCAA
```

FIGURE 372

MGSRGQGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEEQHEKYRPS
QDQGLPASRCLRCCDPGTSMYPATAVPQINITILKGEKGDRGDRGLQGKYGKTGSAGARGHTG
PKGQKGSMGAPGERCKSHYAAFSVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMFTGKFYCYV
PGLYFFSLNVHTWNQKETYLHIMKNEEEVVILFAQVGDRSIMQSQSLMLELREQDQVWVRLYK
GERENAIFSEELDTYITFSGYLVKHATEP

Important features:

Signal sequence.

amino acids 1-25

N-glycosylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

Amidation site.

amino acids 150-154

Cell attachment sequence.

amino acids 104-107

FIGURE 373

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCCG

TTCTAGACGCGGGAAAA**ATG**CTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGGAA

GCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAATGC

ACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCAGAGG

ATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAAAACCCAAAG

ATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAGAGTTCT

TCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATGTGGTTAA

TGATGAGAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACTGGTTCTTCC

TTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTTAAAAAAGGATC

CATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATATGTGGGTATGG

AAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCTCAATATCCCAG

AAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAACAGCTAGCAGTTT

GCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGAAAAGATGTATTTA

ATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCCCAACCAGGTAGTAG

AAGGCTGTTGTTCAGATATGGCTGTTACTTTTAATGGACTGACTCCAAATCAGATGCATGTGA

TGATGTATGGGGTATACCGCCTTAGGGCATTTGGGCATATTTTCAATGATGCATTGGTTTTCT

TACCTCCAAATGGTTCTGACAATGAC**TGA**GAAGTGGTAGAAAGCGTGAATATGATCTTTGTA

TAGGACGTGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAATACAGCTGTATGTTTCTT

TTTCTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAGTCAGTAGTACATTTTTAAA

TGAGGGTGGTTTTTTTCTTTAAAACACATGAACATTGTAAATGTGTTGGAAAGAAGTGTTTTA

AGAATAATAATTTTGCAAATAAACTATTAATAAATATTATATGTGATAAATTCTAAATTATGA

ACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTTAAAAAATTTTAACAGGTCTTTA

GCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATTTGTGATTAAAGTAAAACTTTTAG

CTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCTAAGCCTCCCCAAGTTCCAATGGAT

TTGCCTTCTCAAAATGTACAACTAAGCAACTAAAGAAATTAAAGTGAAAGTTGAAAAAT

FIGURE 374

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERMELSKSFRVYCIILV
KPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRKAYKYAFDKYRDQYNWFFLARPTTFAI
IENLKYFLLKKDPSQPFYLGHTIKSGDLÉYVGMEGGIVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAV
CLKYAGVFAENAEDADGKDVFNTKSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFG
HIFNDALVFLPPNGSDND

Important features:

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 375

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAAT

CTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGAAGAAA

AAAAATC**ATG**AAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCACGGG

GCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCCAA

AGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATTGACAA

CCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGACAAGTG

GTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCGAGATCCA

GAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAACCACCCAAA

GACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATTTCTTCAGATAT

CTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGACCAGAGCCTAC

GGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGACGAATACTTGGA

AATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGACGTGGC

CGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACTATCCACCATACATTTCAGAAGCCAA

GGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCCTCAGCAGTCCCCTC

AGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAAGAAAGGGGTGAAAGT

GGAAAACAGACCTTTCCTCTCAAAACTCATCTTCTTCAATGTCTCTGAACATGACTATGGGAA

CTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGCATCATGTATTTGGTCC

AGGCGCCGTCAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGGCTGCGTCTGGCTGCTGCC

TCTTCTGGTCTTGCACCTGCTTCTCAAATTT**TGA**TGTGAGTGCCACTTCCCCACCCGGGAAAG

GCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACACCGACAGCAACCAATCAGATA

TATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGAAATTTGAGGGAGGGGAACAAA

GAATACTTTGGGGGGAAAGAGTTTTAAAAAGAAATTGAAAATTGCCTTGCAGATATTTAGG

TACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGCACACCCGGCTTGGACCCACTGCA

AGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAAGGGCTCAGCCTCTCTGCCCACAGA

GTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCAAATTCAATCAGTCCATAGAGACGAA

CAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGGGCACTTTGGTAGACTGTGCCACCACG

GCGTGTGTTGTGAAACGTGAAATAAAAGAGCAAAAAAAA

FIGURE 376

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNRV
TRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPKTS
RVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQ
GITREQSGDYECSASNDVAAPVVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAVPSAE
FQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIMLFGPGA
VSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Important features:

Signal peptide:

amino acids 1-28

FIGURE 377

```
CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCTTTAGATTGTGAA
ATGTGGCTCAAGGTCTTCACAACTTTCCTTTCCTTTGCAACAGGTGCTTGCTCGGGGCTGAAG
GTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTCCAC
TATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCACACA
ATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACCAACAC
AAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCTGATGAA
GGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACTCTATCTGCCAGTCAGAAGATA
CAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCTCTGGGGCT
GTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGGCTAGCTTAC
CAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTCTCCCCAAAAC
AATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCTGCCTGGTGAGG
AACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCATCATATATTATGGACCTTATGGA
CTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGTTGACCTTGGAGAG
GCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCCAACACCTACTCCTGGATTAGGAGG
ACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTTGCATCTGAGAAAGTA
GCCCAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAACCGGCAGGCAAGATGAA
ACTCATTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTGCACAGAAAGGAAAATCA
TTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTGATTATATCCATGTGTCTTCTC
TTCCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAAACTAGAAGGCAGGCCAGAA
ACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATGCTCTGGATGACTTCGGAATA
TATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGGATTCCAAGCAGGTCTGTTCCA
GCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGTGTATGAAGTTATTCAGCACATC
CCTGCCCAGCAGCAAGACCATCCAGAGTGAACTTTCATGGGCTAAACAGTACATTCGAGTGAA
ATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAGTATATTAATCTGGAATCAGTGAAGA
AACCAGGACCAACACCTCTTACTCATTATTCCTTTACATGCAGAATAGAGGCATTTATGCAAA
TTGAACTGCAGGTTTTTCAGCATATACAATGTCTTGTGCAACAGAAAAACATGTTGGGGAA
ATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGGGGAGAACGAAGTGACAGGGGTTTCCT
CATAAGTTTTGTATGAAATATCTCTACAAACCTCAATTAGTTCTACTCTACACTTTCACTATC
ATCAACACTGAGACTATCCTGTCTCACCTACAAATGTGGAAACTTTACATTGTTCGATTTTTC
AGCAGACTTTGTTTTATTAAATTTTTATTAGTGTTAAGAATGCTAAATTTATGTTTCAATTTT
ATTTCCAAATTTCTATCTTGTTATTTGTACAACAAAGTAATAAGGATGGTTGTCACAAAACA
AAACTATGCCTTCTCTTTTTTTTCAATCACCAGTAGTATTTTTGAGAAGACTTGTGAACACTT
AAGGAAATGACTATTAAAGTCTTATTTTTATTTTTTTCAAGGAAAGATGGATTCAAATAAATT
ATTCTGTTTTTGCTTTTAAAAAAAAAAAAAA
```

FIGURE 378

MWLKVFTTFLSFATGACSGLKVTVPSHTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPHTMPKYLLGSVNKS
VVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQKIQVTVDDPVTKPVVQIHPPSGAVEY
VGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFSPQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPII
YYGPYGLQVNSDKGLKVGEVFTVDLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMD
YVCCAYNNITGRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQKLEG
RPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRSVPASDCVSGQDLHSTVYEVIQHIPAQQQDHPE

Important features:

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 379

ATAGTAGAAGAATGTCTCTGAAATTACTGGATGAGTTTCAGTCATACTTTCACATGGGCACAA
TTTCACATTCAAGCTCCTTATCCTAGGCTAATTTTATATTATGTTAAATCACTTGTTTTTGTT
CTCACGGCTTCCTGCCTGCTATAGGCATAATTACGAGGAAGCAGAACTTCTCCAGAAGCAAGC
GCACATGCGTTCCAAAATAAGAGCAAATTCGCTCTAAACACAGGAAAAGACCTGAAGCTTTAA
TTAAGGGGTTACATCCAACCCCAGAGCGCTTTTGTGGGCACTGATTGCTCCAGCTTCTGCGTC
ACTGCGCGAGGGAAGAGGGAAGAGGATCCAGGCGTTAGAC**ATG**TATAGACACAAAAACAGCTG
GAGATTGGGCTTAAAATACCCACCAAGCTCCAAAGAAGAGACCCAAGTCCCCAAAACATTGAT
TTCAGGGCTGCCAGGAAGGAAGAGCAGCAGCAGGGTGGGAGAGAAGCTCCAGTCAGCCCACAA
GATGCCATTGTCCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCACCGCTGCCCT
GCCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGGAATAAG
GAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGACCTCCCAGGCCAGTGCCGG
GCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGCACCCCCCCAGCA
ATCCGCGCGCCGGGACAGAATGCCCTGCAGGAACTTCTTCTGGAAGACCTTCTCCTCCTGCAA
A**TAG**

FIGURE 380

MYRHKNSWRLGLKYPPSSKEETQVPKTLISGLPGRKSSSRVGEKLQSAHKMPLSPGLLLLLLS
GATATAALPLEGGPTGRDSEHMQEAAGIRKSSLLTFLAWWFEWTSQASAGPLIGEEAREVARR
QEGAPPQQSARRDRMPCRNFFWKTFSSCK

Important features:

Transmembrane domain:

amino acids 51-69

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 35-39, 92-96

N-myristoylation sites.

amino acids 64-70, 75-81, 90-96

Amidation site.

amino acids 33-37

FIGURE 381

```
GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCGCC
GCGCCGCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCGCGC
CCGCGCCCAGGTGAGCGCTCCGCCCGCCGCGAGGCCCCGCCCCGGCCCGCCCCCGCCCCGCCC
CGGCCGGCGGGGGAACCGGGCGGATTCCTCGCGCGTCAAACCACCTGATCCCATAAAACATTC
ATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCCGCCGCCCTCGCCCTG
TGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGCGGAGCG
CGCCGAGCCTCGTCCCGCGGCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGATGCGGAC
CCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTCAGTCCCCGACGCGCCCCGCCCA
ACCCCTACG**ATG**AAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTGCTGTGGCTG
CAGGCCTGGCAGGTGGCAGCCCCATGCCCAGGTGCCTGCGTATGCTACAATGAGCCCAAGGTG
ACGACAAGCTGCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTGCTGCCAGCCAG
CGCATCTTCCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTCCGTGCCTGCCGC
AACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGCGGCTGCCTTCACT
GGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCCGGTCTGTGGACCCT
GCCACATTCCACGGCCTGGGCCGCCTACACACGCTGCACCTGGACCGCTGCGGCCTGCAGGAG
CTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTACCTGCAGGACAACGCG
CTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCACACACCTCTTCCTGCAC
GGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTGCACAGCCTCGACCGTCTC
CTACTGCACCAGAACCGCGTGGCCCATGTGCACCGCATGCCTTCCGTGACCTTGGCCGCCTC
ATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCACTGAGGCCCTGGCCCCCCTG
CGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTGTGTGACTGCCGGGCACGCCCA
CTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGAGGTGCCCTGCAGCCTCCCGCAA
CGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATGACCTGCAGGGCTGCGCTGTGGCC
ACCGGCCCTTACCATCCATCTGGACCGGCAGGGCCACCGATGAGGAGCCGCTGGGGCTTCCC
AAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGTACTGGAGCCTGGAAGACCAGCTTCG
GCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTGACAGCCCGCCGGGCAACGGCTCTGGC
CCACGGCACATCAATGACTCACCCTTTGGGACTCTGCCTGGCTCTGCTGAGCCCCGCTCACT
GCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTTCCCCACCTCGGGCCCTCGCCGGAGGCCA
GGCTGTTCACGCAAGAACCGCACCCGCAGCCACTGCCGTCTGGGCCAGGCAGGCAGCGGGGGT
GGCGGGACTGGTGACTCAGAAGGCTCAGGTGCCCTACCCAGCCTCACCTGCAGCCTCACCCCC
CTGGGCCTGGCGCTGGTGCTGTGGACAGTGCTTGGGCCCTGC**TGA**CCCCAGCGGACACAAGA
GCGTGCTCAGCAGCCAGGTGTGTGTACATACGGGGTCTCTCTCCACGCCGCCAAGCCAGCCGG
GCGGCCGACCCGTGGGGCAGGCCAGGCCAGGTCCTCCCTGATGGACGCCTGCCGCCCGCCACC
CCCATCTCCACCCCATCATGTTTACAGGGTTCGGCGGCAGCGTTTGTTCCAGAACGCCGCCTC
CCACCCAGATCGCGGTATATAGAGATATGCATTTTATTTTACTTGTGTAAAAATATCGGACGA
CGTGGAATAAAGAGCTCTTTTCTTAAAAAAA
```

FIGURE 382

MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRIF
LHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPATF
HGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLHGNR
ISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAPLRAL
QYLRLNDNPWVCDCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCAVATGP
YHPIWTGRATDEEPLGLPKCCQPDAADKASVLEPGRPASAGNALKGRVPPGDSPPGNGSGPRH
INDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQAGSGGGGT
GDSEGSGALPSLTCSLTPLGLALVLWTVLGPC

Important features:
Signal peptide:
amino acids 1-26

Leucine zipper pattern.
amino acids 135-156

Glycosaminoglycan attachment site.
amino acids 436-439

N-glycosylation site.
amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain
amino acids 411-425

FIGURE 383

TTCGTGACCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTCCTGAACTTGTCTG

AAGCCCTTGTCCGTAAGCCTTGAACTACGTTCTTAAATCTATGAAGTCGAGGGACCTTTCGCTGCTTTTGTAGGG

ACTTCTTTCCTTGCTTCAGCAAC**ATG**AGGCTTTTCTTGTGGAACGCGGTCTTGACTCTGTTCGTCACTTCTTTGA

TTGGGGCTTTGATCCCTGAACCAGAAGTGAAAATTGAAGTTCTCCAGAAGCCATTCATCTGCCATCGCAAGACCA

AAGGAGGGGATTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTCACA

AACATAACAATGGTCAGCCCATTTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGTTGGGACCAGGGCTTGA

AAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGCTCTGGGCTATGGAAAAGAAGGAAAAGGTA

AAATTCCCCCAGAAAGTACACTGATATTTAATATTGATCTCCTGGAGATTCGAAATGGACCAAGATCCCATGAAT

CATTCCAAGAAATGGATCTTAATGATGACTGGAAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGT

TTGAAAAACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTGATAAAGAAGATG

AAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA**TAG**AGATACATCTACCCTT

TTAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAAAGAACATTTTATTTTTATACAATGTTCTTTCT

TGCTTTGTTTTTTATTTTTATATATTTTTTCTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTT

CTTTCTGATAAGTTATTGGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTC

ACAGATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAAATGTTGCAACTGGGAATATACCACGACATGA

GACCAGGTTATAGCACAAATTAGCACCCTATATTTCTGCTTCCCTCTATTTTCTCCAAGTTAGAGGTCAACATTT

GAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCATGTTATAATGAAATAGTTTATGTGTAACTGGCTCTG

AGTCTCTGCTTGAGGACCAGAGGAAAATGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGAT

GTGCAATGCTGAAGTTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGA

GGCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAACCCTATCTCTAC

TAAAAATACAAAGTAGCCCGGCGTGGTGATGCGTGCCTGTAATCCCAGCTACCCAGGAAGGCTGAGGCGGCAGAA

TCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAGATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAGAA

AAGAACACGGTTAATACCATATNAATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGG

CTCCTAGTGATTGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATGTA

TCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGCTAGCGGAATATCCTT

CCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACATTGTATCATAAGATAAAGTAGTAAACCA

GTCTACATTTTCCCATTTCTGTCTCATCAAAAACTGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAG

CACTTTGGGGGCCAAGGAGGGTGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCT

TGTCTCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAGGCTGA

GACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCCACTGCACTCCAGCCTGGG

TGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAGAAGCAGACCTACAGCAGCTACTATTGAATAAATACCTA

TCCTGGATTTT

FIGURE 384

MRLFLWNAVLTLFVTSLIGALIPEPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSLF
HSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPESTLI
FNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVEDIFD
KEDEDKDGFISAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 385

CTCCCACGGTGTCCAGCGCCCAGA**ATG**CGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCTC
CCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGTCC
CTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGTGGG
ATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAATGAAG
GGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGAACCTC
ACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAGTCTTTA
CTGATCTCTCTGTTCGTCTTTCCAGGACCCTGCTGTCCTCCCTCCCCTTCTCCCACCTTCCAG
CCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCCCAGGATTG
ACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGGGCTGAGGCC
CCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACACAGGAACCTCT
CCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCATGCAGCTGGACTCCACC
TCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGGGTGTCCATCCCG
ATGGTCCGCATACTGGCCCCAGTCCTGGTGCTGCTGAGCCTTCTGTCAGCCGCAGGCCTGATC
GCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCACGGAGACACAGAGG
AACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCCCCTTCCCAGGCCCCT
GAGGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGAGCTGGGCTTCTCGAAG
TTTGTCTCAGCG**TAG**GGCAGGAGGCCCTCCTGGCCAGGCCAGCAGTGAAGCAGTATGGCTGGC
TGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAGTCCAGCTGCCCGGACTCC
AGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGCATGTTCCAGCCTGACCTAGAAGCGTTT
GTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTGGAGACTGGGACATCCCTGAT
AGGTTCACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCAGCAGGGCCAGACAAGGCTCAG
TGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGCCTCATGCCCAGTGTCGGACCCT
GCCTTCCTCCCACTCCAGACCCCACCTTGTCTTCCCTCCCTGGCGTCCTCAGACTTAGTCCCA
CGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCTGGGGTGAGACTGGGATTCTGGCTT
CTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCTGTGAAAAACGTGATTCCTGGCCCCA
CCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAGGACTCTGAATTCTAACAATGCCCAGT
GACTGTCGCACTTGAGTTTGAGGGCCAGTGGGCCTGATGAACGCTCACACCCCTTCAGCTTAG
AGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCCCAATAGATCTGCTCTGTCTGCGACACCA
GATCCACGTGGGGACTCCCCTGAGGCCTGCTAAGTCCAGGCCTTGGTCAGGTCAGGTGCACAT
TGCAGGATAAGCCCAGGACCGGCACAGAAGTGGTTGCCTTTNCCATTTGCCCTCCCTGGNCCA
TGCCTTCTTGCCTTTGGAAAAAATGATGAAGAAAACCTTGGCTCCTTCCTTGTCTGGAAAGGG
TTACTTGCCTATGGGTTCTGGTGGCTAGAGAGAAAGTAGAAAACCAGAGTGCACGTAGGTGT
CTAACACAGAGGAGAGTAGGAACAGGGCGGATACCTGAAGGTGACTCCGAGTCCAGCCCCCTG
GAGAAGGGGTCGGGGGTGGTGGTAAAGTAGCACAACTACTATTTTTTTTCTTTTTCCATTATT
ATTGTTTTTTAAGACAGAATCTCGTGCTGCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCA
AACTCCGCCTCCTGGGTTCAAGTGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAG
GCACGCACCACCACACCTGGCTAATTTTTGTACTTTTAGTAGAGATGGGGTTTCACCATGTTG
GCCAGGCTGGTCTTGAACTCCTGACCTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCG
GGATTACAGGCATGAGCCACTGTGTCTGGCCCTATTTCCTTTAAAAAGTGAAATTAAGAGTTG
TTCAGTATGCAAACTTGGAAAGATGGAGGAGAAAAGAAAAGGAAGAAAAAAATGTCACCCA
TAGTCTCACCAGAGACTATCATTATTTCGTTTTGTTGTACTTCCTTCCACTCTTTTCTTCTTC
ACATAATTTGCCGGTGTTCTTTTTACAGAGCAATTATCTTGTATATACAACTTTGTATCCTGC
CTTTTCCACCTTATCGTTCCATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTT
ATAAATAAAATGTTCATCAGCTGCATAAAAAAAAAAAAAA

FIGURE 386

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCSG

TIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDESLLISLFVFP

GPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPGTSQ

YGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRILAPV

LVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGDVISMP

PLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 387

```
GCGCCGGGAGCCCATCTGCCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCGGCACATG
GCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCCAGCTCGCCCGAGGTCCGTCGGAGG
CGCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATCGGG
ATGTCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCACACT
GAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGCTTCCA
GAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAAGTGGTG
ATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCGAGTGGCC
TTTGCTTCCAATTTCCTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGCCCAGTGAT
GAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCATGTCATCTTA
AAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGACAGAAGGAAGT
GACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATTACTGGCAGCGA
ATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATTGACTACAACCAC
CCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTACCAGTGCACAGCA
GGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGTATGTACAAAGCATC
GGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTGATTTTCCTCTTGGTG
TGGCTGCTAATCCGAAGGAAAGACAAAGAAGATATGAGGAAGAAGAGAGACCTAATGAAATT
CGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCTCCTCTTCCTCAGGCTCT
CGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAATAGTGCCTCACGCAGCCAG
CGGACACTGTCAACTGACGCAGCACCCCAGCCAGGGCTGGCCACCCAGGCATACAGCCTAGTG
GGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATGCTAATCTGACCAAAGCAGAA
ACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAAACGGTCTGAATTACAATGGAC
TTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTCTTCTCGTCATTGGAGCTCAAGT
CACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCAGTGAGCATTGCACGGAACAGATT
CAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAAAGGATGTAAGCTGATTCATCTGTA
AAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGGAAAGCAGGAGTCCAAATCTATTTGT
TGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTGAGGTGAATATACCTAAAACTTTTAAT
GTGGGATATTTTGTATCAGTGCTTTGATTCACAATTTTCAAGAGGAAATGGGATGCTGTTTGT
AAATTTTCTATGCATTTCTGCAAACTTATTGGATTATTAGTTATTCAGACAGTCAAGCAGAAC
CCACAGCCTTATTACACCTGTCTACACCATGTACTGAGCTAACCACTTCTAAGAAACTCCAAA
AAAGGAAACATGTGTCTTCTATTCTGACTTAACTTCATTTGTCATAAGGTTTGGATATTAATT
TCAAGGGGAGTTGAAATAGTGGGAGATGGAGAAGAGTGAATGAGTTTCTCCCACTCTATACTA
ATCTCACTATTTGTATTGAGCCCAAAATAACTATGAAAGGAGACAAAAATTTGTGACAAAGGA
TTGTGAAGAGCTTTCCATCTTCATGATGTTATGAGGATTGTTGACAAACATTAGAAATATATA
ATGGAGCAATTGTGGATTTCCCCTCAAATCAGATGCCTCTAAGGACTTTCCTGCTAGATATTT
CTGGAAGGAGAAAATACAACATGTCATTTATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAA
AAGGGATCTAGGAATGCTGAAAGATTACCCAACATACCATTATAGTCTCTTCTTTCTGAGAAA
ATGTGAAACCAGAATTGCAAGACTGGGTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTA
ATATGTCAAGGAAGGTAGCCGGGCATGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAG
GTTGCAGTGAGCCGAGATTATGCCATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC
```

FIGURE 388

MSLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKVV

ITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHVIL

KVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRIDYNH

PGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLIFLLV

WLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANSASRSQ

RTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Important freatures:

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 389

GCGGCACCTGGAAG**ATG**CGCCCATTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCTTCG

CCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCTGT

CCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCCTCAAAGCTCCAGTCCCCA

AAAGGCAAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTACTCA

GCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGGAGAAC

AGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAACTGGGAATGTGA

CAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAGTTTATTC

AGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCACAAGACTGA

ATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAGTAAAGAAATCAGGAACATGAAATTCA

GGTCTAGCTGGGTATTTATTGCAGCAAAAGGCTTGGAACTCCCTTCCGAAATTCAGAGAGAAA

AGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGAGATCCAGATAG

AAGGCTGCATACCCAAAGAACGAAGC**TGA**CACTGCAGGGTCCTGAGTAAATGTGTTCTGTATA

AACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACAGCCCATATTTGAT

GAGTATTTTGGGTTTGTTGTAAACCAATGAACATTTGCTAGTTGTATCAAATCTTGGTACGCA

GTATTTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGAAACTAAAATGAATGG

AAATTCTTAAAAAAAAAA

FIGURE 390

MRPLAGGLLKVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQKC
DHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTATRC
FDMYEGDNSGPMTKFIQSAAPKSLLFMVTYDDGSTRLNNDAKNAIEALGSKEIRNMKFRSSWV
FIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEIQIEGCIPKERS

Important features:

Signal sequence.

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125

FIGURE 391

```
GGGGGCTTTCTTGGGCTTGGCTGCTTGGAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTCGCCGGGAAAGGG
AGGGAAGAAGGAAGGGCGGGGCCGGCCCCCCTGCGCCCGCCCCGCGCCTCTGCGCGCCCCTGTCCGCCCCGGCCC
AGCCCAGCCCAGCCCCGCGGGCCGGTCACACGCGCAGCCAGCCGGCCGCCTCCCGCGCCCAAGCGCGCCGCTCTG
CTGTGCCCTGCGCCCTTGCCCCGCGCCAGCTTCTGCGCCCGCAGCCCGCCCGGCGCCCCCGGTGACCGTGACCCT
GCCCTGGGCGCGGGGCGGAGCAGGC**ATG**TCCCGCCCGGGGACCGCTACCCCAGCGCTGGCCCTGGTGCTCCTGGC
AGTGACCCTGGCCGGGGTCGGAGCCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGATCTGGAG
CCGGGAGCCCTACTACGCGCGCCCGGAGCCCGAGCTCGAGACCTTCTCTCCGCCGCTGCCTGCGGGGCCCGGGGA
GGAGTGGGAGCGGCGCCCGCAGGAGCCCAGGCCGCCCAAGAGGGCCACCAAGCCCAAGAAAGCTCCCAAGAGGGA
GAAGTCGGCTCCGGAGCCGCCTCCACCAGGTAAACACAGCAACAAAAAAGTTATGAGAACCAAGAGCTCTGAGAA
GGCTGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTCAGAGAGAGTTGCCCACCTCTTGGTCTGGA
AACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCACATCGAGGGAG
ACTCAACATCCAGGCGGGCATTAATGAAAATGATTTTTATGACGGAGCGTGGTGCGCGGGAAGAAATGACCTCCA
GCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCATCACTCAAGGGAGGAACTCCCTCTG
GCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGTCACTGTTAAGAATGG
ATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCTACCCGTCCCCATGGT
GGCCCGCTACATCCGCATAAACCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGAGAATGGAGATCCTGGG
CTGCCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACTGATGACCTGGATTTTAA
GCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAATGAAATGTGTCCCAATATCACCAGAATTTA
CAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATCACCCTGGGGAGCATGAAGT
CGGTGAGCCCGAGTTCCACTACATCGCGGGGGCCCACGGCAATGAGGTGCTGGGCCGGGAGCTGCTGCTGCTGCT
GGTGCAGTTCGTGTGTCAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCTGGTGGAGGAGACGCGGATTCA
CGTCCTCCCCTCCCTCAACCCCGATGGCTACGAGAAGGCCTACGAAGGGGGCTCGGAGCTGGGAGGCTGGTCCCT
GGGACGCTGGACCCACGATGGAATTGACATCAACAACAACTTTCCTGATTTAAACACGCTGCTCTGGGAGGCAGA
GGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCCTGAGTGGTTTCTGTCGGAAAATGC
CACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAAAAATCCCTTTTGTGCTGGGCGGCAACCTGCA
GGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGCGGTCCCCCTGGAAGACGCAGGAACACACCCCCAC
CCCCGATGACCACGTGTTCCGCTGGCTGGCCTACTCCTATGCCTCCACACACCGCCTCATGACAGACGCCCGGAG
GAGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGGCACTGTCAATGGGGCCTCCTGGCACACCGTCGCTGG
AAGTCTGAACGATTTCAGCTACCTTCATACAAACTGCTTCGAACTGTCCATCTACGTGGGCTGTGATAAATACCC
ACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACGGGAATCTCTGATCGTGTTCATGGAGCAGGTTCATCG
TGGCATTAAAGGCTTGGTGAGAGATTCACATGGAAAAGGAATCCCAAACGCCATTATCTCCGTAGAAGGCATTAA
CCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTCCTGAACCCTGGAGAGTATGTGGTCACAGCAAA
GGCCGAAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGGCTATGACATGGGGGCCACAAGGTGTGACTTCAC
ACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGGAGAAGTTTGGGAAGCAGCCCGTCAGCCTGCCAGC
CAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGTGGG**TGA**CCCTCCTGGGCCCTTGAGACTCGTCTGGG
ACCCATGCAAATTAAACCAACCTGGTAGTAGCTCCATAGTGGACTCACTCACTGTTGTTTCCTCTGTAATTCAAG
AAGTGCCTGGAAGAGAGGGTGCATTGTGAGGCAGGTCCCAAAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTT
CTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCAGCAGAGAAAAGCTGATGGGAGTGAGAGAACTCAGCAAG
CCAACCTGGGAATCAGAGAGAGAAGGAGAAGGAGGGGAGCCTGTCCGTTCAGAGCCTCTGGCTGCATAGAAAAGG
ATTCTGGTGCTTCCCCTGTTTGCGTGGCAGCAAGGGTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAG
CATTTCCCCAGCTGGGCTGTCCCAAATGTTACCATTTGAGATGCTCCAGGCGTCCTAAGAGAATCCACCCTCTC
TGGCCCTGGGACATTGCAAGCTGCTACAAATAAATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACA
TCAGTGAGCCTCTTGAATCTGTTTAGTCTCCTTTTTTCAACAAAGGAGTGTGTTCAGAAAAGGAGAGAGAGGCTGA
GATCATTCAGGAGTTTGTTGGGCAGCAAGCATGGAGCTTCTTGCACAAATTCTGGGTCCATAAACAACCCCCAAA
GTCCCTGCTGATCCAGTAGCCCTGGAGGTTCCCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCCTGAAGGGCCA
GAAAATTTAGCCTGGATCTCCTCTTTTACCTGCTAGGACTGGAAAGAGCCAGAAGTGGGGTGGCCTGAAGCCCTC
TCTCTGCTTGAGGTATTGCCCCTGTGTGGAATTGAGTGCTCATGGGTTGGCCTCATATCAGCCTGGGAGTTATTT
TTGATATGTAGAATGCCAGATCTTCCAGATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCAT
CAGTTTGGGAAGAATTATTGAATTATCTTGCAAGAAAAAGTATGTCTCACTTTTTGTTAATGTTGCTGCCTCAT
TGACCTGGGAAAAATGAAAAAAAAAAATAAAGCAAATGGTAAGACCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA
```

FIGURE 392

MSRPGTATPALALVLLAVTLAGVGAQGAALEDPDYYGQEIWSREPYYARPEPELETFSPPLPA

GPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHSVR

VAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCAGRN

DLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIFEGNS

EKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTTTDDLD

FKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEFHYIAGA

HGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGGSELGGWS

LGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAAETRAVIAW

MEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSYASTHRLMTDARR

RVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHESQLPEEWENNR

ESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLLNPGEYVVTAKA

EGFTASTKNCMVGYDMGATRCDFTLSKTNMARIREIMEKFGKQPVSLPARRLKLRGRKRRQRG

FIGURE 393

```
GTCCCACATCCTGCTCAACTGGGTCAGGTCCCTCTTAGACCAGCTCTTGTCCATCATTTGCTGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCCTGGCAATTCTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA
TGGCCTTGCCTTGGGGTCCTGCTTGTTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG
AAGGAGCACGGGGCTGATCAAGCCATCCAGGAAACACTGGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTT
TCTGAATCTAGCCCACTTGGCGGTAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
CTACTTATTTCTTTTAGGGGATTGTCAGGAGGTGACCACTCTCACGGTGAAATACCAAGTGTCAGAGGAAGTGCC
ATCTGGTACAGTGATCGGGAAGCTGTCCCAGGAACTGGGCCGGGAGGAGAGGCGGAGGCAAGCTGGGGCCGCCTT
CCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCATTCAGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCG
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCTGGTTTCCTTTGATGTGCTTGCCACAGGGGATTT
GGCTCTGATCCATGTGGAGATCCAAGTGCTGGACATCAATGACCACCAGCCACGGTTTCCCAAAGGCGAGCAGGA
GCTGGAAATCTCTGAGAGCGCCTCTCTGCGAACCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCCC
TAACACCCTGCACACCTACACTCTGTCTCCCAGTGAGCACTTTGCCTTGGATGTCATTGTGGGCCCTGATGAGAC
CAAACATGCAGAACTCATAGTGGTGAAGGAGCTGGACAGGGAAATCCATTCATTTTTTGATCTGGTGTTAACTGC
CTATGACAATGGGAACCCCCCCAAGTCAGGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAG
CCCTGCGTTTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAAACT
GACCGCCACAGACCCTGACCAAGGCCCCAATGGGGAGGTGGAGTTCTTCCTCAGTAAGCACATGCCTCCAGAGGT
GCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCATTCTGCGTCGACCTCTAGACTATGAAAAGAACCC
TGCCTACGAGGTGGATGTTCAGGCAAGGGACCTGGGTCCCAATCCTATCCCAGCCCATTGCAAAGTTCTCATCAA
GGTTCTGGATGTCAATGACAACATCCCAAGCATCCACGTCACATGGGCCTCCCAGCCATCACTGGTGTCAGAAGC
TCTTCCCAAGGACAGTTTTATTGCTCTTGTCATGGCAGATGACTTGGATTCAGGACACAATGGTTTGGTCCACTG
CTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC
CACACTGGACAGAGAGCAGTGGCCCAAATATACCCTCACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGCATTCAGATCAGTGACATCAACGACAATGCACCTGTGTTTGAGAAAAGCAGGTATGA
AGTCTCCACGCGGGAAAACAACTTACCCTCTCTTCACCTCATTACCATCAAGGCTCATGATGCAGACTTGGGCAT
TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACTATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTCAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCCTGTGCTCAGCGATGGAAAAGCCAGCCTCTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCC
CATCGAGACTCCCAATGGCTTGGGCCCAGCGGGCACTGACACACCTCCACTGGCCACTCACAGCTCCCGGCCATT
CCTTTTGACAACCATTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCCTCTACAGCATCCGCAATGG
AAATGAAGCCCACCTCTTCATCCTCAACCCTCATACGGGGCAGCTGTTCGTCAATGTCACCAATGCCAGCAGCCT
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCCCTTACAGACCCGAGCCCTGTT
GAGGGTCATGTTTGTCACCAGTGTGGACCACCTGAGGGACTCAGCCCGCCAAGCCTGGGGCCTTGAGCATGTCGAT
GCTGACGGTGATCTGCCTGGCTGTACTGTTGGGCATCTTCGGGTTGATCCTGGCTTTGTTCATGTCCATCTGCCG
GACAGAAAAGAAGGACAACAGGGCCTACAACTGTCGGGAGGCCGAGTCCACCTACCGCCAGCAGCCAAGAGGCC
CCAGAAACACATTCAGAAGGCAGACATCCACCTCGTGCCTGTGCTCAGGGGTCAGGCAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCACAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTT
CCACCTCACCCCGACCCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGT
GCTGCAAGACACGGTCAACCTCCTTTTCAACCATCCCAGGCAGAGGAATGCCTCCCGGGAGAACCTGAACCTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCACGTTCCAGGCCTCTGAAGGTTGCAGGCAGCCCCACAGGGAGGCTGGC
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACCAGCCTCCTCTGCAACCCTGAGACGGCAGCGACATCT
CAATGGCAAAGTGTCCCCTGAGAAGAATCAGGGCCCCGTCAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCCTTCGCCGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTCAGCAAATCTCCCAGCTGCT
GTCCTTGCTGCATCAGGGCCAATTCCAGCCCAAACCAAACCACCGAGGAAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGGAGTGCAATCCCAGACACAGATGGCCCAAGTGCAAGGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCCTCAC
CACCAACTACCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTCGG
CAAGGCAGGCACCAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCCGCCTCCGAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCCTCAGTTTAGACTTGGCCACCAGTGCAGCCTCAGGCATGAAAGTGCAAGGGACCCAGGTGG
AAAGACGGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGGTGCCTGTGAACATACCTCAGACGCCT
CTGGATCCAAGAACCAGGGGCCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAAATCTTGTAACTCACTAGCTAG
CGGCGGCCTGAGAACTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC
TGACCAAAGCAGCCCCTTGTAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGTT
TCCTGGCAAAACATATGTGGAGCACAAAGGGTCAGTCCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGGTGGCCTTCTTGGGTAGCAGGAGTCAGGGGGCTGTACCCTGGGGGTGCCAGGAAATGCTCTCTGACCTAT
CAATAAAGGAAAAGCAGTAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 394

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQAG

AAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQVL

DINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIVGPD

ETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESSLALE

IQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPLDYEKN

PAYEVDVQARDLGPNPIPAHCKVLIKVLDVNDNIPSIHVTWASQPSLVSEALPKDSFIALVMA

DDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQDQGLQPLS

AKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVSYRIQDSPV

AHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDANDNAPEVVQ

PVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLLTTIVARDADSGA

NGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGSPPLQTRALLRV

MFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEKKDNRAYNCREAE

STYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEAGWDPCLQAPFHLT

PTLYRTLRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPEPQPATGQPRSRPLKV

AGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQILRSLVRLSVAAFAE

RNPVEELTVDSPPVQQISQLLSLLHQGQFQPKPNHRGNKYLAKPGGSRSAIPDTDGPSARAGG

QTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPDPAWMARLSLPLTTNYRD

NVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSLLEMLLEQRSSMPVEAASE

ALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTEGKSRGSSSSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 amd 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 395

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAGG

CTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAATC

AGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACC**ATG**GGACGCCCCCGACCTCGTGCG

GCCAAGACGTGGATGTTCCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGCACAG

GAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGGCCTTG

TTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTTACAGCT

GCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGAT

GGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGAT

GTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCC

AAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCA

GGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTGTGCAGAAGTA

AAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATGGCATGGTC

TGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGCGATTCTGGAGGCCCCCTGGTGTGT

GATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCT

GGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATCATAGGCAGCAAGGGC

TGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCTCTGGTTC

FIGURE 396

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVLV
GGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLLQL
RDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCEDAYP
GQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWI
KKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 397

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCC**ATG**TCGGACCTGCTAC

TACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGGGT

ACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCACTG

TGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGCTGCA

GCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCCCCCTG

ATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCCCTGAGC

TCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGCCATGTGG

TGACAGCCACCTTCCCCTACACCACCATTCTGTCCATCTGGCTGGCTACCCGCCGTGTCCATC

CTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGGAGATCTACC

AGGAAGACCAGATCCATTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTATGTGCCTGAGA

TGAAGGAGACAGAGTGGAAATGGCGGGGGCTTGTGGAGGCCATTGACACCCAGGTGGATGGCA

CAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCCCTGGCAGCCGGG

AGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGATGACGGTGACACCC

GCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGAGGAGCTGGACTTGG

AGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGCCCCTGGGGACTACCA

AGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAG**TAA**CCCATGGCCTGCACCCTCC

TGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCTCCAGCCCTCTTCCTCCT

TCCTCTGGGGGAGGAGGGGTTCCTGAGGGACCTGACTTCCCCTGCTCCAGGCCTCTTGCTAAG

CCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCAGGGACTATTTTCTGCACCA

GCCCCCAGGGCTGCCGCCCCTGTTGTGTCTTTTTTTCAGACTCACAGTGGAGCTTCCAGGACC

CAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAAAAAAAAAAAAA

FIGURE 398

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGRL
FTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFSFP
APSHVVTATFPYTTILSIWLATRRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPLARQGD
FYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAATLSPGASSRGW
DDGDTRSEHSYSESGASGSSFEELDLEGEGPLGESRLDPGTEPLGTTKWLWEPTAPEKGKE

FIGURE 399

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGGC

TCCGTGCCGCCAAGTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCGAG

AGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTGTTTTTTAAAACT

TCTGTTTCTTGGGAGGGGGTGTGGCGGGGCAGG**ATG**AGCAACTCCGTTCCTCTGCTCTGTTTC

TGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGACGGCTG

GAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGGTTTAAC

CTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTCGGCCACAGCCAGCCC

TTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTCATCATTCACGGATGGACGATG

AGCGGTATCTTTGAAAACTGGCTGCACAAACTCGTGTCAGCCCTGCACACAAGAGAGAAAGAC

GCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGATGCGGTCAAT

AATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGGAGAAGGACGAT

TTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTGGCCGGGTATGCA

GGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGCCGGGCCCATGTTT

GAAGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTGTGGÁTGTCCTCCAC

ACCTACACGCGTTCCTTCGGCTTGAGCATTGGTATTCAGATGCCTGTGGGCCACATTGACATC

TACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGTCTTGGGATCAATTGCA

TATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCGTCCACCTCTTTGTTGAC

TCTCTGGTGAATCAGGACAAGCCGAGTTTTGCCTTCCAGTGCACTGACTCCAATCGCTTCAAA

AAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCATTGGCTACAATGCCAAGAAA

ATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGGCAGGCATGCCTTTCAGAGGT

AACCTTCAGTCCCTGGAGTGTCCC**TGA**GGAAGGCCCTTAATACCTCCTTCTTAATACCATGCT

GCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCACAATCCAATCAAATCGTTGCAA

ATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTTACAAAATAAACAGTGTGGACCCC

TAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 400

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHEG
CYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWLPL
AHQLYTDAVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTVGRI
TGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDFQPGC
GLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLVNQDKPSFAFQCTDSNRFKKGICLSCRKNR
CNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:
Signal peptide:
amino acids 1-16

Lipases, serine active site.
amino acids 163-172

N-glycosylation sites.
amino acids 80-83 and 136-139

FIGURE 401

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCTG
CCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCTGT
CAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGGCTC
AACTCTCCTGCACGCTCAGCCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGGTACC
AGCAGCGGGCAGGCAGTGCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCACCACC
GGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCACAATGCCTGTGTCC
TCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGCTACGGCT
TTAGTCCC**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCTCTGCCTCCCATTTCTGCCCCTGA
CCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGGTTAATAATA
TTCAACATGTCAACAAC

FIGURE 402

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAGS
APRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVQPEDDADYYCSVGYGFSP

FIGURE 403

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGAAAGCAGCGAGTTG

GCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCTCACAACAAG**ATG**CTCAAGGTGTCAGC

CGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCAGTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGG

GCGGTCGGACGGCGGTAATTTTCTGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGG

ACAGTGGAACAAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCCTTCGATCA

GGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTATGCATTGCTCAAGATTC

TCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATGAAAGAAGCAGGAGTAGACCATAGGCAGTG

GAGGGGTCCCATATTATCCACCTGCAAGCAGTGCCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCA

TACCTACTCTTTTCAGTGCAAACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGG

ACATTGCCCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCTGGAGTT

CAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAAGTCAAAACAAGAAGACAAA

AACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATCTTGCCAATTTGCAAGGACTCACTTGGCTGGAT

GTTTAACAGACTTGATACAAACTATGACCTGCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAA

TGAACAGTGTACCAAGGCATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTG

CTACTGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCAAGGGGTAAA

GAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGCCAACACAATGTCATGGCAGTGT

TGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTCATGGGATCCAGAATAAATGGTGTTGCAGATTGTGC

TATAGATTTTGAGATCTCCGGAGATTTTGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGA

CGATATTATGAATGATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGA

CCATGATGTATACATT**TGA**TTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTACAAAAATGATAG

CCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCACATATATTTTGTATAATTATTTGAA

AAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAATAAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACA

AAAAGAAAATACATATGCAGTCTAGTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTCACGAGA

ACAAACTTTGTAAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAGAT

AATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGGAAAAATATGCATGCT

TTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAGGATAACAGAGAGATACCACATGACTCCA

AAAAAAAAAAAAAA

FIGURE 404

MLKVSAVLCVCAAAWCSQSLAAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFRD

EVEDDYFRTWSPGKPFDQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEAGV

DHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPCPSDK

PTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLLRPERSRFDTSILPICKDS

LGWMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQRQQDP

PCQTELSNIQKRQGVKKLLGQYIPLCDEDGYYKPTQCHGSVGQCWCVDRYGNEVMGSRINGVA

DCAIDFEISGDFASGDFHEWTDDEDDEDDIMNDEDEIEDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 405

```
GGAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCCAGACCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGACTTCCAAGAGTGACTCCGTCGGAGGAAAATGACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCCTGCTGAGTCTGCTCTTCCTGGTCCAAGGTGCCCACGGCAGGGGCCACAGG
GAAGACTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG
CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCCTTTCCCTGCAGCCCACCCTGCTTCCCGA
TCCTTCCCTGACCCCAGGGGCCTCTACCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTC
TATGGCAAGCGTGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTG
GCTCAGGGCCCCCCGCTGTTAGCCACTTCTGTCACCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCCAGTGCC
GCCAGCTTCACCTTCTCCTTCCACAGTCCTCCCCACACGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTCATGGGGGACATGGTGTCCTTCGAG
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCACAGCCGGCCTCCAGGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCTCGAACACTCTTCCAGAGGACGAAAGGC
CGGAGCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCGTG
GTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTTGAAGACCCCACA
TTGAGCAGCCCGGGGCATTGGAGCAGTGCTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGC
AACCACTTGACCTACTTTGCAGTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGC
CTCCTCTCCTACGTGGGCTGTGTCGTCTCTGCCCTGGCCTGCCTTGTCACCATTGCCGCCTACCTCTGCTCCAGG
GTGCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTC
CTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCC
ATCTTCCTGCACTTCTCCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTG
GTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAG
GGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGC
CTGGTGTTTCTGTTCAACATGGCCATGCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCCACACCCAA
AAGTGGTCACATGTGCTGACACTGCTGGGCCTCAGCCTGGTCCTTGGCCTGCCCTGGGCCTTGATCTTCTTCTCC
TTTGCTTCTGGCACCTTCCAGCTTGTCGTCCTCTACCTTTTCAGCATCATCACCTCCTTCCAAGGCTTCCTCATC
TTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGGCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCC
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCCATGTGATGAAG
CAGAGATGCGGCCTCGTCGCACACTGCCTGTGGCCCCCGAGCCAGGCCCAGCCCCAGGCCAGTCAGCCGCAGACT
TTGGAAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGACTCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGGACTACTCGGCTCTCACTCAGCTCCCACGGGACTCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCCCTGGGCCCAGCCCTCATTGCTGGG
GGCCAGGCCTTGGATCTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCCTGCCTGGGACAGAAATGTGGCTCCA
GTTGCTCTGTCTCTCGTGGTCACCCTGAGGGCACTCTGCATCCTCTGTCATTTTAACCTCAGGTGGCACCCAGGG
CGAATGGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC
AGCTCGCCTACCTCTGAGCCCAGGCCCCCTCCCTCCCTCAGCCCCCCAGTCCTCCCTCCATCTTCCCTGGGGTTC
TCCTCCTCTCCCAGGGCCTCCTTGCTCCTTCGTTCACAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTGGGGA
GTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCCTCGGCCTGCCCCTGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTG
CTCACCCTGACCAAGCACACGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGA
CCATGCCAGTCCCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCTA
GAGCCTGACACTCTCCTAAGAGGTTCTCTCCAAGCCCCCAAATAGCTCCAGGCGCCCTCGGCCGCCCATCATGGT
TAATTCTGTCCAACAAACACACACGGGTAGATTGCTGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACTCCTCCTGCCAACATTCAGTCTGGTATGTGAGGCGTGCGTGAAGCAAGAACTCCTGGAGCTACAGGGACA
GGGAGCCATCATTCCTGCCTGGGAATCCTGGAAGACTTCCTGCAGGAGTCAGCGTTCAATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAAAAGAAGTACATGTTCATTGTAGAGA
ATTTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAATCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 406

MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSEE
ALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQHQE
ESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDLQLLSQF
LKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQDLHIH
SRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGEKVLGI
VVQNTKVANLTEPVVLTFQHQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRETQTSCF
CNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPLPCRRKPRDY
TIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLEGYNLYRLV
VEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIYPSMCWIRDS
LVSYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLGLPWALIFFSF
ASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLPISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590 and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327 and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 407

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGAG

CCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTCGC

TCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCTGGA

CTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAAGCCC

TGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAAAGCTC

AGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAGCAGGGC

TCTCAGAAGGCGGTGGTGCCCAGCTGGGATC**ATG**TTGTTGGCCCTGGTCTGTCTGCTCAGCTG

CCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTGCTACATGA

CTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGCTTATTTCAC

AAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACAACGGGATCTT

CCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAACGTGTGCCGGAT

GTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGCCATGAAGATAAC

CCAAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCCAGGGAAAAGACCT

CACTGAATGGGTGGATGGCTGTGACTTC**TAG**GATGGACGGAACCATGCACAGCAGGCTGGGAA

ATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATAAAGGATGGTTGAACG

TGAAA

FIGURE 408

MLLALVCLLSCLLPSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALDY
EADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYCSDLLNPNLKDTVICAMKITQEPQGLGYWE
AWRHHCQGKDLTEWVDGCDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homolgous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 409

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
CTCTTTCAGCCCGGGATCGCCCCAGCAGGG**ATG**GGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
GCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGCCGGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCTT
CCCGCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTA
GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA
TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTCAGCACCATT
TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
AGACTAAGCAAAAGTGGGCACATACAAATTCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGTCAGCCATTCAAGTTTAT
ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACT**TAA**AACTCCAAACTAGAGTACGTAACATTGAAA
AATGAGGCATAAAAATGCAATAAACTGTTACAGTCAAGACCATTAATGGTCTTCTCCAAAATATTTTGAGATATA
AAAGTAGGAAACAGGTATAATTTTAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
TTTCCTAACTTTGAAAAATTTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC
AGTCTGTTTTTAACAGGTTCTATTACCCAGAACTTTTTTGTAAATGCGGCAGTTACAAATTAACTGTGGAAGTTT
TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCAA
CTTTTCTCTATTTACATATGCATCTCTCCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
AGATTTTTATAACCAAATACATTTCAGTGTAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
CCAAAAGCTGACATTTTCACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
AAATGAAGAATATAGTTTAAAAGCTTCCTCCTCCATAGGGACACATTTTCTCTAACCCTTAACTAAAGTGTAGGA
TTTTAAAATTAAATGTGAGGTAAAATAAGTTTATTTTTAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
TAATCATGTTATGTTAATTTTAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA
TTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
AGAATTTAGAAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACTATAAATAAATATCTAGA
ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AAACTTTGGCTGTAGGTTTTTATTTTCTACAAGAATTCTGGTTTGAATTATTTTTGTAAGCAGGTACATTTTATA
AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAAAAATTTTTATTATAACAACTTTTAT
TAAAATGGCCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTAAA
CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC
TCAAGTACTAGTAATTTAACTTCATCATGAATGAACTATAATTTTTAAGTTATGCCCATTTATAACGTTGTTTAT
GACTACATTGTGAGTTAGAAACAAACTTAAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT
CTTGATGAGCAATAATGATAACCAGAGAGTGATTTCATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
TCTTAGGCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
ATGATCAATTACCTTAATTGGCCAAGCAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
AGGTCAATAAGATCCTTGCCTATGAATACCCCTCCCTTTTGCGCTGTTAAATTTGCAATGAGAAGCAAATTTACA
GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAAACTGTGATTAAGAATTCTA
CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAATGAATTTGTTACATAATCTTCT
ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACTTCTTACCATATAAAAACGATAATTGCTT
TATTTGGAAAAGAATTTAGGAATACTAAGGACAATTATTTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
CATAACCAAAAAAGCAAAACTTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAAGTTATGCTAAATTGAGTAAGCTGTTTATCACTT
AACAGCTCATTTTGTCTTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTC
CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
GTTATGCTGTTCTTTCATGTGAATGTCAAGACATGGAGGGTGTTTGTAATTTTATGGTAAAATTAATCCTTCTTA
CACATAATGGTGTCTTAAAATTGACAAAAAATGAGCACTTACAATTGTATGTCTCCTCAAATGAAGATTCTTTAT
GTGAAATTTTAAAAGACATTGATTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG
CTCAAACTGCTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
AAATTATCAAAGGAAAA

FIGURE 410

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEYQ

VLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCFDNTFSTISEKVIFFELIL

DNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRNIQE

SNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 411

```
CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAAACTTCCCAGGGGACCGCATTCCAGAGTCA
GTGACTCTGTGAAGCACCCACATCTACCTCTTGCCACGTTCCCACGGGCTTGGGGGAAAGATGGTGGGGACCAAG
GCCTGGGTGTTCTCCTTCCTGGTCCTGGAAGTCACATCTGTGTTGGGGAGACAGACGATGCTCACCCAGTCAGTA
AGAAGAGTCCAGCCTGGGAAGAAGAACCCCAGCATCTTTGCCAAGCCTGCCGACACCCTGGAGAGCCCTGGTGAG
TGGACAACATGGTTCAACATCGACTACCCAGGCGGGAAGGGCGACTATGAGCGGCTGGACGCCATTCGCTTCTAC
TATGGGGACCGTGTATGTGCCCGTCCCCTGCGGCTAGAGGCTCGGACCACTGACTGGACACCTGCGGGCAGCACT
GGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTCTGGTGCCTCAACAGGGAGCAGCGGCCTGGCCAGAACTGC
TCTAATTACACCGTACGCTTCCTCTGCCCACCAGGATCCCTGCGCCGAGACACAGAGCGCATCTGGAGCCCATGG
TCTCCCTGGAGCAAGTGCTCAGCTGCCTGTGGTCAGACTGGGGTCCAGACTCGCACACGCATTTGCTTGGCAGAG
ATGGTGTCGCTGTGCAGTGAGGCCAGCGAAGAGGGTCAGCACTGCATGGGCCAGGACTGTACAGCCTGTGACCTG
ACCTGCCCAATGGGCCAGGTGAATGCTGACTGTGATGCCTGCATGTGCCAGGACTTCATGCTTCATGGGGCTGTC
TCCCTTCCCGGAGGTGCCCCAGCCTCAGGGGCTGCTATCTACCTCCTGACCAAGACGCCGAAGCTGCTGACCCAG
ACAGACAGTGATGGGAGATTCCGAATCCCTGGCTTGTGCCCTGATGGCAAAAGCATCCTGAAGATCACAAAGGTC
AAGTTTGCCCCCATTGTACTCACAATGCCCAAGACTAGCCTGAAGGCAGCCACCATCAAGGCAGAGTTTGTGAGG
GCAGAGACTCCATACATGGTGATGAACCCTGAGACAAAAGCACGGAGAGCTGGGCAGAGCGTGTCTCTGTGCTGT
AAGGCCACAGGGAAGCCCAGGCCAGACAAGTATTTTTGGTATCATAATGACACATTGCTGGATCCTTCCCTCTAC
AAGCATGAGAGCAAGCTGGTGCTGAGGAAACTGCAGCACCAGGCTGGGGAGTACTTTTGCAAGGCCCAGAGT
GATGCTGGGGCTGTGAAGTCCAAGGTTGCCCAGCTGATTGTCACAGCATCTGATGAGACTCCTTGCAACCCAGTT
CCTGAGAGCTATCTTATCCGGCTGCCCCATGATTGCTTTCAGAATGCCACCAACTCCTTCTACTATGACGTGGGA
CGCTGCCCTGTTAAGACTTGTGCAGGGCAGCAGGATAATGGGATCAGGTGCCGTGATGCTGTGCAGAACTGCTGT
GGCATCTCCAAGACAGAGGAAAGGGGAGATCCAGTGCAGTGGCTACACGCTACCCACCAAGGTGGCCAAGGAGTGC
AGCTGCCAGCGGTGTACGGAAACTCGGAGCATCGTGCGGGGCCGTGTCAGTGCTGCTGACAATGGGGAGCCCATG
CGCTTTGGCCATGTGTACATGGGGAACAGCCGTGTAAGCATGACTGGCTACAAGGGCACTTTCACCCTCCATGTC
CCCCAGGACACTGAGAGGCTGGTGCTCACATTTGTGGACAGGCTGCAGAAGTTTGTCAACACCACCAAAGTGCTA
CCTTTCAACAAGAAGGGGAGTGCCGTGTTCCATGAAATCAAGATGCTTCGTCGGAAAGAGCCCATCACTTTGGAA
GCCATGGAGACCAACATCATCCCCCTGGGGGAAGTGGTTGGTGAAGACCCCATGGCTGAACTGGAGATTCCATCC
AGGAGTTTCTACAGGCAGAATGGGGAGCCCTACATAGGAAAAGTGAAGGCCAGTGTGACCTTCCTGGATCCCCGG
AATATTTCCACAGCCACAGCTGCCCAGACTGACCTGAACTTCATCAATGACGAAGGAGACACTTTCCCCCTTCGG
ACGTATGGCATGTTCTCTGTGGACTTCAGAGATGAGGTCACCTCAGCAGCCACTTAATGCTGGCAAAGTGAAGGTC
CACCTTGACTCGACCCAGGTCAAGATGCCCAGAGCACATATCCACAGTGAAACTCTGGTCACTCAATCCAGACACA
GGGCTGTGGGAGGAGGAAGGTGATTTCAAATTTGAAAATCAAAGGAGGAACAAAAGAGAAGACAGAACCTTCCTG
GTGGGCAACCTGGAGATTCGTGAGAGGAGGCTCTTTAACCTGGATGTTCCTGAAAGCAGGCGGTGCTTTGTTAAG
GTGAGGGCCTACCGGAGTGAGAGGTTCTTGCCTAGTGAGCAGATCCAGGGGGTTGTGATCTCCGTGATTAACCTG
GAGCCTAGAACTGGCTTCTTGTCCAACCCTAGGGCCTGGGGCCGCTTTGACAGTGTCATCACAGGCCCCAACGGG
GCCTGTGTGCCTGCCTTCTGTGATGACCAGTCCCCTGATGCCTACTCTGCCTATGTCTTGGCAAGCCTGGCTGGG
GAGGAACTGCAAGCAGTGGAGTCTTCTCCTAAATTCAACCCAAATGCAATTGGCGTCCCTCAGCCCTATCTCAAC
AAGCTCAACTACCGTCGGACGGACCATGAGGATCCACGGGTTAAAAAGACAGCTTTCCAGATTAGCATGGCCAAG
CCAAGGCCCAACTCAGCTGAGGAGAGCAATGGGCCCATCTATGCCTTTGAGAACCTCCGGGCATGTGAAGAGGCA
CCACCCAGTGCAGCCCACTTCCGGTTCTACCAGATTGAGGGGGATCGATATGACTACAACACAGTCCCCTTCAAC
GAAGATGACCCTATGAGCTGGACTGAAGACTATCTGGCATGGTGGCCAAAGCCGATGGAATTCAGGGCCTGCTAT
ATCAAGGTGAAGATTGTGGGGCCACTGGAAGTGAATGTGCGATCCCGCAACATGGGGGGCACTCATCGGCGGACA
GTGGGGAAGCTGTATGGAATCCGAGATGTGAGGAGCACTCGGGACAGGGACCAGCCCAATGTCTCAGCTGCCTGT
CTGGAGTTCAAGTGCAGTGGGATGCTCTATGATCAGGACCGTGTGGACCGCACCCTGGTGAAGGTCATCCCCCAG
GGCAGCTGCCGTCGAGCCAGTGTGAACCCCATGCTGCATGAGTACCTGGTCAACCACTTGCCACTTGCAGTCAAC
AACGACACCAGTGAGTACACCCATGCTGGCACCCTTGGACCCCACTGGGCCACAACTATGGCATCTACACTGTCACT
GACCAGGACCCTCGCACGGCCAAGGAGATCGCGCTCGGCCGGTGCTTTGATGGCACATCCGATGGCTCCTCCAGA
ATCATGAAGAGCAATGTGGGAGTAGCCCTCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTC
CAGTACCTCCAAAGCACCCCAGCCCAGTCCCCTGCTGCAGGCACTGTCCAAGGAAGAGTGCCCTCGAGGAGGCAG
CAGCGAGCGAGCAGGGGTGGCCAGCGCCAGGGTGGAGTGGTTGGCCTCTCTGAGATTTCCTAGAGTTGCTCAACAG
CCCCTGATCAACTAAGTTTTGTGGTACTTCACCCTCTTCTGCCCTCATTTCATGTGACAGCCATTGTGAGACTGA
TGCACAAACTGTCACTTGGTTAATTTAAGCACTTCTGTTTTCGTGAATTTGCTTGTTTGTTTCTTCATGCCTTTA
CTTACTTTGTCCCATGCTACTGATTGGCACGTGGCCCCCACAATGGCACAATAAAGCCCCTTTGTGAAACTGTTC
TTTAAATGAAACACAAGAAATTGGCCACTGGTAAAACTCTGCAGCTTCAACTGTACTTCATTTAATGCCATTAAT
GCAAATATACTTCCTCTTCTTTTTGCATGGTTTTGCCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAA
ATAACCAATATAAAGCATATTTCTTGGCCTTGCTCCACAGGACATAGGCAAGCCTTGATCATAGTTCATACATAT
AAATGGTGGTGAAATAAAGAAATAAAACACAATACTTTTACTTGAAATGTAAATAACTTATTTATTTCTTTGCTA
AATTTGGAATTCTAGTGCACATTCAAAGTTAAGCTATTAAATATAGGGTGATCATAGTTCCTCTACCAAGTCTGG
AAAGAACATCTCCTGGTATCCACAATTACACCAGGTTGCTAACTGTATTTGTACATTTCCCTTTGCATTCGCTTT
TGTTCTTGCTAGAAACCCAGTGTAGCCCAGGGCAGATGTCAATAAATGCATACTCTGTATTTCGAAAAAA
```

FIGURE 412

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNID
YPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQRP
GQNCSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCSEAS
EEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTKTPKL
LTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTSLKAATIKAEFVRAETPYMVMNP
ETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAGEYFCKA
QSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQNATNSFYYDVGRCPVKTCAGQQ
DNGIRCRDAVQNCCGISKTEEREIQCSGYTLPTKVAKECSCQRCTETRSIVRGRVSAADNGEP
MRFGHVYMGNSRVSMTGYKGTFTLHVPQDTERLVLTFVDRLQKFVNTTKVLPFNKKGSAVFHE
IKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRQNGEPYIGKVKASVTFLDPR
NISTATAAQTDLNFINDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHLDSTQVKMPEHI
STVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNLDVPESRRCFVKV
RAYRSERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGACVPAFCDDQSPDA
YSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLNYRRTDHEDPRVKKTAFQISMAKPR
PNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFNEDDPMSWTEDYLAWW
PKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRSTRDRDQPNVSAACLEF
KCSGMLYDQDRVDRTLVKVIPQGSCRRASVNPMLHEYLVNHLPLAVNNDTSEYTMLAPLDPLG
HNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALTFNCVERQVGRQSAFQYLQ
STPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVAQQPLIN

FIGURE 413

GCCACGTTGTCTTCTTTCCTTCACCACCACCCAGGAGCTCAGAGATCTAAGCTGCTTTCCATC

TTTTCTCCCAGCCCCAGGACACTGACTCTGTACAGG**ATG**GGGCCGTCCTCTTGCCTCCTTCTC

ATCCTAATCCCCCTTCTCCAGCTGATCAACCCGGGGAGTACTCAGTGTTCCTTAGACTCCGTT

ATGGATAAGAAGATCAAGGATGTTCTCAACAGTCTAGAGTACAGTCCCTCTCCTATAAGCAAG

AAGCTCTCGTGTGCTAGTGTCAAAAGCCAAGGCAGACCGTCCTCCTGCCCTGCTGGGATGGCT

GTCACTGGCTGTGCTTGTGGCTATGGCTGTGGTTCGTGGGATGTTCAGCTGGAAACCACCTGC

CACTGCCAGTGCAGTGTGGTGGACTGGACCACTGCCCGCTGCTGCCACCTGACC**TGA**CAGGGA

GGAGGCTGAGAACTCAGTTTTGTGACCATGACAGTAATGAAACCAGGGTCCCAACCAAGAAAT

CTAACTCAAACGTCCCACTTCATTTGTTCCATTCCTGATTCTTGGGTAATAAAGACAAACTTT

GTACCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 414

MGPSSCLLLILIPLLQLINPGSTQCSLDSVMDKKIKDVLNSLEYSPSPISKKLSCASVKS

QGRPSSCPAGMAVTGCACGYGCGSWDVQLETTCHCQCSVVDWTTARCCHLT

FIGURE 415

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCGGTGTG
AGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAGGA
GTGTGTGGAACAGGACCCGGGACAGAGGAACC**ATG**GCTCCGCAGAACCTGAGCACCTTTTGCC
TGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTGGGGG
TGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAACTAGCCCTGCAGCTTC
ATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTGCTGCTT
ATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAAGGATTAA
AAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTCACACTTCTTTGGGGATTTTGGTTTCA
TGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATATTATTGTAG
ATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTTAGAAACAAAC
CTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGATGCGGACCACCC
AGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAATGCCCTAATGTCA
AACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTGAGAGACGGCATGG
AGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGGAGATTTACGGTTCC
GAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATTTGTACACAAATGTGA
CAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACTCACTTGGATGGTCACA
AGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCTATGGAAGAAAGGGGAAG
GGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAATCACTTTTGATGTGGATT
TTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAACAGCTACTGAAACAAGGGT
CAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TGA**GAGTGAATAAAATTGGACTTTGTTT
AAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTTTTGTGTGTGTTTTTGTTTTTA
TTTTCAATATGCAAGTTAGGCTTAATTTTTTTATCTAATGATCATCATGAAATGAATAAGAGG
GCTTAAGAATTTGTCCATTTGCATTCGGAAAAGAATGACCAGCAAAAGGTTTACTAATACCTC
TCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGTTTCAAGAATTAAAGCTGCAAGAGG
ACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGAGTTGTTAGCAATTTCATTCAAAATG
CCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTTGTTATTTTTA

FIGURE 416

MAPQNLSTFCLLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQAQEKFQDLGAAYE
VLSDSEKRKQYDTYGEEGLKDGHQSSHGDIFSHFFGDFGFMFGGTPRQQDRNIPRGSDIIVDLEVTLEEVYAGNF
VEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQMTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMEYPFI
GEGEPHVDGEPGDLRFRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLW
KKGEGLPNFDNNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY   ·

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 417

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGGA

TGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCGCCCGGGGCAG

G**ATG**ACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTCATGATCCTGCT

GATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCTCTAG

GCCGCACACGGGGCCGCCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACGGCCGA

CTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGACCTTCC

CAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAGGCTACGA

CTGGTCCCCGCGCGACGCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAGCGGAGGAG

CGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCGCGCATTCGA

CGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGGCCATCTACTG

CTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTGAGCGGAAGCCT

GCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCACGTGCACAACGC

CAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCTCCCGCCACCTCAT

GAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCCTTCGTGCGCCTGAT

CTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCGCAAGTTCGCCGTGCC

CATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCGCCTCGGCGCGCGAGGCCTTCCGCGC

TGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGACCCGCACACGGAGAAGCT

GGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCACCCGTGCCAGATCGACTA

CGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGCAGCTGCTGCAGCTACTCCA

GGTGGACCGGCAGCTCCGCTTCCCCCCGAGCTACCGGAACAGGACCGCCAGCAGCTGGGAGGA

GGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCTGTATAAACTCTACGAGGCCGA

CTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCCGAGAC**TGA**AAGCTTTCGCGTTG

CTTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCCAGTTTTTTTATGACCTACGATTT

TGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCCATTGAGTACTGTATCGATATTGTT

TTTTAAGATTAATATATTTCAGGTATTTAATACGA

FIGURE 418

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRDRELTAD
SDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAERRS
VLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLSGSL
LHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPFVRLI
SAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDPHTEKL
APFNEHWRQVYRLCHPCQIDYDFVGKLETLDEDAAQLLQLLQVDRQLRFPPSYRNRTASSWEE
DWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 419

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAAG

GCTGCCAGGAAGGAGACGCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGACT

GTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGCTGA

AGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATCATCA

ACACCATTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAAGATCA

ACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGTCGGGCA

CGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAATGCCATCG

TGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGAACGCTTTG

GGCTGTTAGGGGGCTCCAAGGTCCTGGCCAAGAAAGAGCTGGCCTATGTCCCAATTATCGGCT

GGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGATCGCAAGACGG

TTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCCTGATTCACTGTG

AGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCCGGGCCAAGGGGC

TGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATCACCGTGAGGAGCT

TGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAATAATGAAAATCCAA

CACTGCTGGGAGTCCTAAACGGAAAGAAATACCATGCAGATTTGTATGTTAGGAGGATCCCAC

TGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCACAAGCTCTACCAGGAGA

AGGATGCCTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGAGACGCCCATGGTGCCCC

CCCGGCGGCCCTGGACCCTCGTGAACTGGCTGTTTTGGGCCTCGCTGGTGCTCTACCCTTTCT

TCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACGCTGGCCAGCTTCATCCTCG

TCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGTGACGGAAATTGACAAGGGCT

CTGCCTACGGCAACTCTGACAGCAAGCAGAAACTGAATGAC**TGA**CTCAGGGAGGTGTCACCAT

CCGAAGGGAACCTTGGGGAACTGGTGGCCTCTGCATATCCTCCTTAGTGGGACACGGTGACAA

AGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGACCTCTCCAGCCAGGGAGTCTGGT

CTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTTTTTTCCCCATGTGCTTTAGTGGGC

TTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGCTGTGTGGTGAGTGTGAACTTTGTTC

TGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAGGGCAGGGCTGGGGACCGAAGGGGACA

AGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGTAACCCTTGGTTGCCAGAGATAAAGTGA

AAAGTGCTTTAGGTGAGATGACTAAATTATGCCTCCAAGAAAAAAAATTAAAGTGCTTTTCT

GGGTCAAAAAAAAAAAA

FIGURE 420

MDLAGLLKSQFLCHLVFCYVFIASGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLVM

LLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHKFEIDFLCGWSLSERFGLLGGSKVLAKKEL

AYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQHLRDYPEKYFFLIHCEGTRFTEKKHEISM

QVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKKYHAD

LYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVNWLFWA

SLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDSKQKLND

FIGURE 421

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCATC

GCC**ATG**GACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGCCC

TGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTCACC

ACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGCGGCG

CTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGGGTGCC

CTGAAGGAGGAGGTCGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTGCAGACC

ACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCTGCGGGAA

CTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACGTCCGCACT

GAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCGTGCCCCACG

TCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGACGTGGGCGGCG

GCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCCTGGATGAGCAG

GGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGGGCTGTGCGCCAT

CTGGGCAAGGTTCAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTTCAGCCACTGGAAC

CAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTCATGATGCTGCACACGGGGCTG

TGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAGAAAAGGCACAACTGC

TGACCCCGCCCAGTGCCCTGGAGCCGCGCCCATTGCAGCATGTCGTATCCTGGGGGCTGCTCA

CCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCTCATCCACCGCTGCTGAG

TCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCCTGGGCTCTGGGACCTCCA

TGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACCTCCACTAGCTCCAAAATCC

CTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAACCAAGGTTAGGTGACTGAGG

ACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACTGGAAGCTGTTTTTGCAGCCTGAGG

AAGCATCAATAAATATTTGAGAAATGAAAAAA

FIGURE 422

MDTTRYSKWGGSSEEVPGGPWGRWVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAAL
LDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALREL
RERVTQGLAEAGRGREDVRTELFRALEAVRLQNNSCEPCPTSWLSFEGSCYFFSVPKTTWAAA
QDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFSHWNQ
GEPNDAWGRENCVMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 423

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGGC
GAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCGGC
GCCCAAC**ATG**GCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGATCGC
GGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGGTCCA
GCCCATGACCGCCTCCAACTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTTTACGC
CCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGCTTTTGCAAAGAATGGTGA
AATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTGGCCGCTT
CTTTGTCACCACTCTCCCAGCATTTTTTCATGCAAAGGATGGGATATTCCGCCGTTATCGTGG
CCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATCAGTCGAGCC
TCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTCTTTTTAGCAT
CTCTGGCAAGATATGGCATCTTCACAACTATTTCACAGTGACTCTTGGAATTCCTGCTTGGTG
TTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCTTTTTATGGGTCTGGTCTTGGT
GGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGCGTTCTGAGCAGAA
TCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAGGAGGAAAAAGATGA
TTCAAATGAAGAAGAAAACAAAGACAGCCTTGTAGATGATGAAGAAGAGAAAGAAGATCTTGG
CGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACTTGGCTGCTGGTGTGGATGAGGA
GAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTGACCCGGGAGGAAGTAGA
GCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGCTGACACAGAGGTGGTGGA
AGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGACTG**TAG**ATTTAATGATGCGT
TTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGGCCTGCAGTTTGTACCAAATC
CTTAATTTTTCCTGAATGAGCAAGCTTCTCTTAAAAGATGCTCTCTAGTCATTTGGTCTCATG
GCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGTGACAATCAGGATATAGAAAAAC
AAACGTAGTGTTGGGATCTGTTTGGAGACTGGGATGGGAACAAGTTCATTTACTTAGGGGTCA
GAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATCAGCACCTTCCAGAGACAAGGCTGC
AGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCTCCTGAGCATCCCCAAAGTGTAACGT
AGAAGCCTTGCATCCTTTTCTTGTGTAAAGTATTTATTTTTGTCAAATTGCAGGAAACATCAG
GCACCACAGTGCATGAAAAATCTTTCACAGCTAGAAATTGAAAGGGCCTTGGGTATAGAGAGC
AGCTCAGAAGTCATCCAGCCCTCTGAATCTCCTGTGCTATGTTTTATTTCTTACCTTTAATT
TTTCCAGCATTTCCACCATGGGCATTCAGGCTCTCCACACTCTTCACTATTATCTCTTGGTCA
GAGGACTCCAATAACAGCCAGGTTTACATGAACTGTGTTTGTTCATTCTGACCTAAGGGGTTT
AGATAATCAGTAACCATAACCCCTGAAGCTGTGACTGCCAAACATCTCAAATGAAATGTTGTG
GCCATCAGAGACTCAAAAGGAAGTAAGGATTTTACAAGACAGATTAAAAAAAAATTGTTTTGT
CCAAAATATAGTTGTTGTTGATTTTTTTTTAAGTTTTCTAAGCAATATTTTTCAAGCCAGAAG
TCCTCTAAGTCTTGCCAGTACAAGGTAGTCTTGTGAAGAAAGTTGAATACTGTTTTGTTTTC
ATCTCAAGGGGTTCCCTGGGTCTTGAACTACTTTAATAATAACTAAAAAACCACTTCTGATTT
TCCTTCAGTGATGTGCTTTTGGTGAAAGAATTAATGAACTCCAGTACCTGAAAGTGAAAGATT
TGATTTTGTTTCCATCTTCTGTAATCTTCCAAGAATTATATCTTTGTAAATCTCTCAATACT
CAATCTACTGTAAGTACCCAGGGAGGCTAATTTCTTT

FIGURE 424

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAPW
CPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRGPG
IFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAWCSY
VFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEKDDSN
EEENKDSLVDDEEEKEDLGDEDEAEEEEEEDNLAAGVDEERSEANDQGPPGEDGVTREEVEPE
EAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

Important features:
Signal peptide:
amino acids 1-22

Transmembrane domain:
amino acids 191-211

N-glycosylation site.
amino acids 46-49

Thioredoxin family proteins.   (homologous region to disulfide isomerase)
amino acids 56-72

Flavodoxin proteins
amino acids 173-187

FIGURE 425

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAATCCCGTGCGCCGCGG
CTGGGCCGTCGGAGAGTGCGTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTCACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTCACCCT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTCAC
TTAAATCAGAACTTGCATAAGAAAGAGA**ATG**GGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTTACAGTTTACTTGG
AGTGTCCAAAACTGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACATTGGAAAGAAGAGAATTTGATGC
TGCTGTTAATTCTGGAGAACTGTGGTTTGTAAATTTTTACTCCCCAGGCTGTTCACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTTACTTCGAATTGGAGCTGTTAACTGTGGTGATGATAGAAT
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTCCATACAAACTGCTTTTGCTGCTGGTATTGGCTGGCTGATCACTTTTTGTTCAAA
AGGAGGAGATTGTTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCCAGATTTTGAACTACTTTCGGCAAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTTATTTTTTCATTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAACT
AAAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGACTGTTCCTCTGCACCAGACATCTGTAGTAA
TCTGTATGTTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAATATGAAATTCATCATGGAAA
GAAGATTCTATATGATATACTTGCCTTTGCCAAAGAAAGTGTGAATTCTCATGTTACCACGCTTGGACCTCAAAA
TTTTCCTGCCAATGACAAAGAACCATGGCTTGTTGATTTCTTTGCCCCTGGTGTCCACCATGTCGAGCTTTACT
ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTCATGA
GGGACTCTGTAACATGTATAACATTCAGGCTTATCCAACAACAGTGGTATTCAACCAGTCCAACATTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACCTTCAACGAACTAGTTACACAAAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
GTGTCATCCTTGCCAAGTCTTAATGCCAGAATGGAAAAGAATGGCCCGGACATTAACTGGACTGATCAACGTGGG
CAGTATAGATTGCCAACAGTATCATTCTTTTTGTGCCCAGGAAAACGTTCAAAGATACCCTGAGATAAGATTTTT
TCCCCCAAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGATGCTTATTCCCTGAGAAT
CTGGGGTCTAGGATTTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTACAAGG
GAAAAATCATTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGCTCCAGAATTTGAGCT
CTTGGCTAGGATGATTAAAGGAAAAGTGAAAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCCTATCCAACTGTTAAGTTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAGCAATCGCTGCCTTAATAAGTGAAAAATTGGAAACTCTCCGAAATCAAGGCAA
GAGGAATAAGGATGAACTT**TGA**TAATGTTGAAGATGAAGAAAAAGTTTAAAAGAAATTCTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCA
GAATTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAACTTTTTCTGTAAAGGGCCGGTTTATAAATATTTTA
GACTTTGCAGGCTATAATATATGGTTCACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAACAACCTTTAAAAAAATATTAAAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCAGTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCACGTTTTTTGGCTGACCTGAAAAGAGGTAACT
TAGTTTTTGGTCACTTGTTCTCCTAAAAATGCTATCCCTAACCATATATTTATATTTCGTTTTAAAAACACCCAT
GATGTGGCACAGTAAACAAACCCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTTTTCTTTCCTTCTCA
AAGGTTGAAAAATGCTTTTAATTTTTCACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATTCTGTAGTTTGCTGTATCATCCAGGAAAACCTGAGGGAAAAAATTA
TAGCAATTAACTGGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTCATGTATTTTCTGAAATTGCTTTCATAGAAATTTTCCCACTGATAGTTGATTTTTGAGGCATCTAATAT
TTACATATTTGCCTTCTGAACTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTCATAGTTTTGG
TTTTTCACTCCTGTCCAGTCTATTTATTATTCAAATAGGAAAAATTACTTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACTGTAGTTATTCCAGTTACTAGTTTACTGTCAGAGGGCTGCCTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTTATAAGAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTAGA
CTCAAAGAATCACAAATTTGTCAGTAACATGTAGTTGTTTAGTTATAATTCAGAGTGTACAGAATGGTAAAAATT
CCAATCAGTCAAAAGAGGTCAATGAATTAAAAGGCTTGCAACTTTTTCAAAAAAAAAAAAAAAAAA

FIGURE 426

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKLH

PDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYYRYDFGIYD

DDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNCGDD

RMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNSIQTA

FAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSANTLEDR

LAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQPSLAVFK

GQGTKEYEIHHGKKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWCPPCRALL

PELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIHEYEGHHSAEQILEFI

EDLMNPSVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMARTLTGLINVGS

IDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLGFLPQVSTDLT

PQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPEFELLARMIKGKVKAGKVDCQAYAQTCQKAG

IRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 427

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGCA

CGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTTGA

GTCCTTCTGAG**ATG**ATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTCGCGATGGTAG

CGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTCAATT

CCAACGCTATCAAGAACCTGCCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTCTGCAG

TCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACAACTACC

AGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGTCCCACCC

GCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACGCTGCATGC

GTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTTCTGATCAAA

ATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATAGCACCT

TGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAAAGGACAAGAAG

GTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTAGACACTTCTGGT

CCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCATAGGAGAAAAGGCT

CTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTCTTGCCGGATACAGA

AAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGAGACAC**TAA**ACCAGCT

ATCCAAATGCAGTGAACTCCTTTTATATAATAGATGCTATGAAAACCTTTTATGACCTTCATC

AACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCATTCCAATAACACCTTCCA

AAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACTCCCTGTGATTGCAGTAAATTACT

GTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGAAACTTTTAATTATTTTTCT

AAAGGTGCTGCACTGCCTATTTTTCCTCTTGTTATGTAAATTTTTGTACACATTGATTGTTAT

CTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATTTCAGCTTATAGTTCTTAAAAG

CATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCAAGGATCTCTTGGAATGACAAAT

GATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATTTTCTGAAATGTACTATCTTAATG

CTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGAAATAAAATTTAACATTTAAAAAAA

AAAAAA

FIGURE 428

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSAA
PGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRHAM
CCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEGSVC
LRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQKDHH
QASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 429

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCCC

TTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGACCC

CAGCGTTACC**ATG**CATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCT

CCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGAATAT

AGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGTCGTTT

CAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATTTCCAAA

TGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATA

CAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATA

CAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAGTGACCCCAT

TCAAGAAATTCGGGACTTAGCAGAAATCACCACTCTTGATCGCAGCAAAAGAAATATCATTGG

ATATTTTGAGCAAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCGAATATTTTGCA

TGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAAACCGGAAAGATATAGTGGCGA

CAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACTTGGGAGCTATGAC

AAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCTCTTGTCCGAGAAATAAC

ATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCATACTCTTTCACATGAA

AGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGCAATTAATAAGTGAAAA

AGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACATCCTCTTCTGCACATACA

GAAAACTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAGGCATATGTATGTGTTTGG

AGACTTCAAAGATGTATTAATTCCTGGAAAACTCAAGCAATTCGTATTTGACTTACATTCTGG

AAAACTGCACAGAGAATTCCATCATGGACCTGACCCAACTGATACAGCCCCAGGAGAGCAAGC

CCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAAACTAGCACCCAGTGAATATAG

GTATACTCTATTGAGGGATCGAGATGAGCTT**TAA**AAACTTGAAAACAGTTTGTAAGCCTTTC

AACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTATATTTTCATAATTCTATGTGTAT

TTTTATTTTGAATAAACAGAAAGAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAA

FIGURE 430

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQM
LHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYRGQ
RSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFERVANILHDDC
AFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREITFEN
GEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLHIQKTP
ADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPGEQAQDV
ASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:
Signal peptide:
amino acids 1-29

Endoplasmic reticulum targeting sequence.
amino acids 403-406

Tyrosine kinase phosphorylation site.
amino acids 203-211

Thioredoxin family proteins
amino acids 50-66

FIGURE 431

GAGCAGGACGGAGCC**ATG**GACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGCA

GGCTGGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCGTG

CAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGCGTG

GACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGCAGTG

CGGGGTTGCGGTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGCTTCTG

GCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTCACCTCG

CGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTGCTACAGC

TGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCTGCTACAAC

GCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCAGCTAATGTG

ACTGTGTCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGATGGAGTAACA

GGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACTCTGACCTCCGC

AACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCTCCAGAGCCCACG

ACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAGACCCACATCCACC

ACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAGAACACGAGGCCTCC

CGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGACCGCAGCAATTCAGGG

CAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTGTGTGGCTCCCACAGCT

GGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCCTACT**GTGA**GCTTCTCCACCTGGA

AATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCTCATCACTTCCTGTTCCCA

CCACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCCCAGTATCCCCAGCTTCTGC

TGCGCTGGTTTGCGGCTTTGGGAAATAAAATACCGTTGTATATATTCTGCCAGGGGTGTTCTA

GCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTCTCCGCTTGTCCTCTTGTGATG

TTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTGAGAGAGAGGATGCTAAGCTTCC

TACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGGGGTGGGTGGGACAATGGCTCCCC

ACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGGGAATCGGTTCCCCATATGTCTTCC

TTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTACCCAATTCGCCCTATAGTGAGTCGTA

FIGURE 432

MDPARKAGAQAMIWTAGWLLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCTE
AVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRALDP
AGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTVSLP
VRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPPLVRLPPPEPTTVAST
TSVTTSTSAPVRPTSTTKPMPAPTSQTPRQGVEHEASRDEEPRLTGGAAGHQDRSNSGQYPAK
GGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

FIGURE 433

CGGGACTCGGCGGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC**ATG**GAGTT
GGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAGGA
GAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTTCGC
TGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTTTCAA
TCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCCAATGC
AACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAACCTGAG
GCGGCTGCTTGAACCTTTGGATGCAAATGTCGATGCT**TAA**GAAAACCGGCCACTTCAGCAACA
GCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCCACCCTATCCCCTCTAACACCATT
CCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCCTGCCCACCTC
CCGTGATGTGTGTGTGTGTGTGTGTGTGTGACTGTGTGTGTTTGCTAACTGTGGTCTTTGTGG
CTACTTGTTTGTGGATGGTATTGTGTTTGTTAGTGAACTGTGGACTCGCTTTCCCAGGCAGGG
GCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCCGTGGCCCTCCATCACCTTCTGCTCCTA
GGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGCGTAGGGTAAGAGC
ACGGGCAGTGGTCTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCACTTTGTCATCATTCT
TCATGGACTCCTTTCACTCCTTTAACAAAAACCTTGCTTCCTTATCCCACCTGATCCCAGTCT
GAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCCCAGCGTTGACGTCAGG
CAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACGAGGAGTCCCCATCTGCC
CCGCCCCTTCACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCTACTCTGCCCCTGGGGAAT
GTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGGGACCCTACCCCTTCCAACC
TTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAGATGCAGACTACAGTCCCTGC
AATTGGGTCTCTGGCAGGCAATAGTTGAAGGACTCCTGTTCCGTTGGGGCCAGCACACCGGGA
TGGATGGAGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACGTCCCCTTAGATGGGCAGCAGAG
GCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGAGCGGTGAGCGAGGTGGGTTGGAG
ACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAGGTTGAAGGTCATAACGAGAGTGGG
AACTCAACCCAGATCCCGCCCCTCCTGTCCTCTGTGTTCCCGCGGAAACCAACCAAACCGTGC
GCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTATCCTCAACAACAACAGAAAAAAGGAAT
AAAATATCCTTTGTTTCCT

FIGURE 434

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRRCKC
SFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSLWNLRRLLEPLDANVDA

FIGURE 435

GGTCCTTA**ATG**GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTGC

TGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGTCA

TCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGACTT

TTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAACTAA

ATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACTTACAG

AGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCTGCAGGCAA

GGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGTTTCGATG

GGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCCTGGAGCCA

GAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATTACTTCTCAA

TGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACCCTGGAGCCAA

GTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCACAGCCACCACCC

TCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCATCTGAGGAGAGT

CCTTTAGAGTGACAGGTTAAAGCTGATACCAAAAGGCTCCTGTGAGCACGGTCTTGATCAAAC

TCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGTGGCCTCCAGCAGAT

CATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTTTTGCCAACAATTTTA

CCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACCTGATGGAATTCCTGCA

CTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTCTTTTTGTTTGGAAAA

TCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATTGTCAGTAAAATAATCACG

TTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAGAATTTTTAAATTATTTAAT

AAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTATTGTTCTGTACTGATATTTAA

ATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAA

FIGURE 436

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFLH
YDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQARMS
CEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFSMGD
CIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

FIGURE 437

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC**ATG**AAGCTCTTGTGGCAGG
TAACTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTCGTCTACCTCACGGCGCAAGTGTGGATTCTGT
GTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCCAGAACTGCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCA
GCAAGGTGGTGTGCACGCGCCGGGGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACC
TCATGGAGAACAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCCTGCAGTTGG
GCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAACACCCTGGAGCTGTTCG
ACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTCCAAGCTGCGGGAGCTCTGGCTTCGCAACA
ACCCCATCGAAAGCATCCCCTCTTACGCCTTCAACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCA
AGAAGCTGGAGTATATCTCTGAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCA
ACATTAAAGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACTTCCCTG
AGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCATGAACTCACAGGTCAGCCTGA
TTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAACTCAACTTGGCCCACAATAACCTCTCTTCTTTGC
CCCATGACCTCTTTACCCCGCTGAGGTACCTGGTGGAGTTGCATCTACACCACAACCCTTGGAACTGTGATTGTG
ACATTCTGTGGCTAGCCTGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCATGCTC
CCATGCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCTTCATCATGG
ACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGTCGGACTCCCCCTATGTCCTCCG
TGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGCCTCCCGCCACCCAAGGATCTCTGTCCTCAACGACG
GCACCTTGAACTTTTCCCACGTGCTGCTTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCA
ACTCCAACGCCTCGGCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACTACAGCTTCTTCACCACAG
TAACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCCTACCACGTCCACTG
GTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACCCGTGTGCCCAAGCAGGTGGCAGTAC
CCGCGACAGACACCACTGACAAGATGCAGACCAGCCTGGATGAAGTCATGAAGACCACCAAGATCATCATTGGCT
GCTTTGTGGCAGTGACTCTGCTAGCTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGC
GGAGTACAGTCACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGCAG
CAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCACAATTCATGACCATATTAACT
ACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGCCTGGGGAACTCTCTGCACCCCACAGTCA
CCACTATCTCTGAACCTTATATAATTCAGACCCATACCAAGGACAAGGTACAGGAAACTCAAATA**TGA**CTCCCCT
CCCCCAAAAAACTTATAAAATGCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTT
TTCTTGTATATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAAATTTAAAGACAAAA
AGTCAAAACA

FIGURE 438

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCTR

RGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLASLN

TLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEYISE

GAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEELEMSGNHFPEIRPGSFHGLSSLKKLWVMNS

QVSLIERNAFDGLÅSLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHNPWNCDCDILWLAWWL

REYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAELKCRTPP

MSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSNASAYLNV

STAELNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQTTRVPKQV

AVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRSTVTAARTVE

IIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTENSLGNSLHPT

VTTISEPYIIQTHTKDKVQETQI

FIGURE 439

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACC**ATG**AGGCTG

TCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCTGC

CCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTCCAA

GTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTGCACC

GATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAA**TAG**TGAAAAAAT

GTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACCCTGATC

TTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

FIGURE 440

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEVK
HCTDQISFKKRLSLKKSWWK

FIGURE 441

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGATG

GGGTTGCTGGTTTAAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCACC

ACCTCCGCCAGGAACTGCAGGCCCACCTGTCTGCAACCCAGCTGAGGCC**ATG**CCCTCCCCAGG

GACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCTCCAG

CTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCAGCCAA

GCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCGGAAGATGGAGGTCAAGCAGAAGGGGC

AGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGTCAGGGGT

TCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGGGAAGAGGC

CAAAGAGGCCCCAGCCGACAAG**TGA**TCGCCCACAAGCCTTACTCACCTCTCTCTAAGTTTAGA

AGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTACAAGCTCAGG

AGGCGAATAAATGTTCAAACTGTA

FIGURE 442

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDGG
QAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

FIGURE 443

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTCG

TGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAA**ATG**AACACGTGGCTGCTGTTC

CTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTGCTC

CTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCTGAGT

ATGTCCCCACCCTAAGCCCCGATCCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATCTTACA

CCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCACGCCTGGGGCCAGAGTCTTTGTCCCC

CGTGTGCGCATGTGTTCAGGGTCAGCCTCTCCCAGAAGTGAGATCATGGACAAAAAGGGCAAA

TCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCAGCAAGAAGCTGAACTCACGCCG

AGACCTGCAGGAGTGGTGCCAGGTGCT**TGA**AGTAACAAGTTTAAAATGTTCAGAGACAATGGA

ATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTTCCAAAAACAC

AAGTAGAAATTCTAACAATGAAATATATTACAGGCAGGTCACCCACTAACCAAACAACTGAAG

CGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTCAGTCATGTTGCT

GAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAACCTGTGTTGGAAGTAACAACAACCTCC

CTGCTCCTGGCACCAGCCGTTTTGGTCATGGTGGGCCAGCTGCAAAGCGTCTTCCATTCTCTG

GGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGACACGGGCAGCAGAGTG

TGTCCAGGCCAGCCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCAACCCCTGGTCAGGGCA

GAGGGAGTTGGGTGGGTCAGGCTCTGGGCTCACCTCCATCTCCAGAGCATCCCCTGCCTGCAG

TTGTGGCAAGAACGCCCAGCTCAGAATGAACACACCCCACCAAGAGCCTCCTTGTTCATAACC

ACAGGTTACCCTACAAACCACTGTCCCCACACAACCCTGGGGATGTTTAAAACACACACCTC

TAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGAATTTTTTTTAATGAAAGTGC

AATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAA

FIGURE 444

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDFLGLVHLGQLLIFHIYLSMSPTLSPRSPQGWV
VRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQAQ
QEAELTPRPAGVVPGA

FIGURE 445

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGA**ATG**GACTGGCCTCACAACCTGC

TGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAAGA

GGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACCTGG

TGTCACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATGGTGG

CCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCTGTGGA

TGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCCGTATCC

CCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACCATGCAGG

AGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCTGTGCGCCGCCGCCTCTGCC

CGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCGCTGTGGGCT

GCACCTGCATCTTC**TGA**ATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGACCATCCTCCT

TGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAG

FIGURE 446

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEYE
RNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCLGC
VNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 447

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC**ATG**GCCAAGA
TGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCTAT
CACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGGTGC
CCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTGGATG
GAGATACCAACACATCCACCCAGGAGGTGGTACAATACAACTGGGAGACTGGGGATGACCGGT
TCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAAACTGTGGAAGAACCAG
GGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAAGGACTAC
TGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAAGCGGTTGA
TGGAGAAGGCTTCCCTCCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTATGGTTATCC
CTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCTACTAACAGAC
TTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACTGAGCGCCTTTGCTGCTGTTTCCTCT
GTCCTGTCAGGTCTCCTGGGGATGGTGGCCCACATGATGTATTCACAAGTCTTCCAAGCGACT
GTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTGGGCCTTCTACATG
GCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCAACACGTACACCAGG
ATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAACCCGAACTGCCTACCA
CATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCACCGTGGGTCCTTTGACC
AGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGGGAGTCGACTTCTACTCC
GAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTGAAAGAAGCAGTTAGGTCA
TCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGAGTAGGCTTGAGCCCTACCT
TACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCGTCTCTTGAGCATGGTTTTTA
GAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTCCTAAGGGATTCCTGGGTGCCA
CTGCTCTCTTTTCCTCTACAGCTCCATCTTGTTTCACCCACCCCACATCTCACACATCCAGAA
TTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGCTAAACCATGGAGATAAAAGAAG
AGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

FIGURE 448

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMPV
SLDGDTNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKRGE
KGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQLPPA
TNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLETTCLELWLG
LLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

FIGURE 449

```
CCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCC
ACGCGTCCGGTGCAAGCTCGCGCCGCACACTGCCTGGTGGAGGGAAGGAGCCCGGGCGCCTCTCGCCGCTCCCCG
CGCCGCCGTCCGCACCTCCCCACCGCCCGCCGCCCGCCGCCCGCCGCCCGCAAAGCATGAGTGAGCCCGCTCTCT
GCAGCTGCCCGGGGCGCGAATGGCAGGCTGTTTCCGCGGAGTAAAAGGTGGCGCCGGTCAGTGGTCGTTTCCAAT
GACGGACATTAACCAGACTGTCAGATCCTGGGGAGTCGCGAGCCCCGAGTTTGGAGTTTTTTCCCCCCACAACGT
CACAGTCCGAACTGCAGAGGGAAAGGAAGGCGGCAGGAAGGCGAAGCTCGGGCTCCGGCACGTAGTTGGGAAACT
TGCGGGTCCTAGAAGTCGCCTCCCCGCCTTGCCGGCCGCCCTTGCAGCCCCGAGCCGAGCAGCAAAGTGAGACAT
TGTGCGCCTGCCAGATCCGCCGGCCGCGGACCGGGGCTGCCTCGGAAACACAGAGGGGTCTTCTCTCGCCCTGCA
TATAATTAGCCTGCACACAAAGGGAGCAGCTGAATGGAGGTTGTCACTCTCTGGAAAAGGATTTCTGACCGAGCG
CTTCCAATGGACATTCTCCAGTCTCTCTGGAAAGATTCTCGCTA**ATG**GATTTCCTGCTGCTCGGTCTCTGTCTAT
ACTGGCTGCTGAGGAGGCCCTCGGGGGTGGTCTTGTGTCTGCTGGGGGCCTGCTTTCAGATGCTGCCCGCCGCCC
CCAGCGGGTGCCCGCAGCTGTGCCGGTGCGAGGGGCGGCTGCTGTACTGCGAGGCGCTCAACCTCACCGAGGCGC
CCCACAACCTGTCCGGCCTGCTGGGCTTGTCCCTGCGCTACAACAGCCTCTCGGAGCTGCGCGCCGGCCAGTTCA
CGGGGTTAATGCAGCTCACGTGGCTCTATCTGGATCACAATCACATCTGCTCCGTGCAGGGGGACGCCTTTCAGA
AACTGCGCCGAGTTAAGGAACTCACGCTGAGTTCCAACCAGATCACCCAACTGCCCAACACCACCTTCCGGCCCA
TGCCCAACCTGCGCAGCGTGGACCTCTCGTACAACAAGCTGCAGGCGCTCGCGCCCGACCTCTTCCACGGGCTGC
GGAAGCTCACCACGCTGCATATGCGGGCCAACGCCATCCAGTTTGTGCCCGTGCGCATCTTCCAGGACTGCCGTA
GCCTCAAGTTTCTCGACATCGGATACAATCAGCTCAAGAGTCTGGCGCGCAACTCTTTCGCCGGCTTGTTTAAGC
TCACCGAGCTGCACCTCGAGCACAACGACTTGGTCAAGGTGAACTTCGCCCACTTCCCGCGCCTCATCTCCCTGC
ACTCGCTCTGCCTGCGGAGGAACAAGGTGGCCATTGTGGTCAGCTCGCTGGACTGGGTTTGGAACCTGGAGAAAA
TGGACTTGTCGGGCAACGAGATCGAGTACATGGAGCCCCATGTGTTCGAGACCGTGCCGCACCTGCAGTCCCTGC
AGCTGGACTCCAACCGCCTCACCTACATCGAGCCCCGGATCCTCAACTCTTGGAAGTCCCTGACAAGCATCACCC
TGGCCGGGAACCTGTGGGATTGCGGGCGCAACGTGTGTGCCCTAGCCTCGTGGCTCAGCAACTTCCAGGGGCGCT
ACGATGGCAACTTGCAGTGCGCCAGCCCGGAGTACGCACAGGGCGAGGACGTCCTGGACGCCGTGTACGCCTTCC
ACCTGTGCGAGGATGGGGCCGAGCCCACCAGCGGCCACCTGCTCTCGGCCGTCACCAACCGCAGTGATCTGGGGC
CCCCTGCCAGCTCGGCCACCACGCTCGCGGACGGCGGGGAGGGGCAGCACGACGGCACATTCGAGCCTGCCACCG
TGGCTCTTCCAGGCGGCGAGCACGCCGAGAACGCCGTGCAGATCCACAAGGTGGTCACGGGCACCATGGCCCTCA
TCTTCTCCTTCCTCATCGTGGTCCTGGTGCTCTACGTGTCCTGGAAGTGTTTCCCAGCCAGCCTCAGGCAGCTCA
GACAGTGCTTTGTCACGCAGCGCAGGAAGCAAAAGCAGAAACAGACCATGCATCAGATGGCTGCCATGTCTGCCC
AGGAATACTACGTTGATTACAAACCGAACCACATTGAGGGAGCCCTGGTGATCATCAACCGAGTATGGCTCGTGTA
CCTGCCACCAGCAGCCCGCGAGGGAATGCGAGGTG**TGA**TTGTCCCAGTGGCTCTCAACCCATGCGCTACCAAATA
CGCCTGGGCAGCCGGGACGGGCCGGCGGGCACCAGGCTGGGGTCTCCTTGTCTGTGCTCTGATATGCTCCTTGAC
TGAAACTTTAAGGGGATCTCTCCCAGAGACTTGACATTTTAGCTTTATTGTGTCTTAAAAACAAAAGCGAATTAA
AACACAACAAAAAACCCCACCCCACAACCTTCAGGACAGTCAGTCTATCTTAAATTTCATATGAGAACTCCTTCCTCCC
TTTGAAGATCTGTCCATATTCAGGAATCTGAGAGTGTAAAAAAGGTGGCCATAAGACAGAGAGAGAATAATCGTG
CTTTGTTTTATGCTACTCCTCCCACCCTGCCCATGATTAAACATCATGTATGTAGAAGATCTTAAGTCCATACGC
ATTTCATGAAGAACCATTGGAAAGAGGAATCTGCAATCTGGGAGCTTAAGAGCAAATGATGACCATAGAAAGCTA
TGTTCTTACTTTGTGTGTGTGTCTGTATGTTTCTGCGTTGTGTGTCTTTGTAGGCAAGCAAACGTTGTCTACACA
AACGGGAATTTAGCTCACATCATTTCATGCCCCTGTGCCTCTAGCTCTGGAGATTGGTGGGGGGAGGTGGGGGGA
AACGGCAGGAATAAGGGAAAGTGGTAGTTTTAACTAAGGTTTTGTAACACTTGAAATCTTTTCTTTCTCAAATTA
ATTATCTTTAAGCTTCAAGAAACTTGCTCTGACCCCTCTAAGCAAACTACTAAGCATTTAAAAGAGAATCTAATT
TTTAAAGGTGTAGCACCTTTTTTTTTTATTCTTCCCACAGAGGGTGCTAATCTCATTATGCTGTGCTATCTGAAAA
GAACTTAAGGCCACAATTCACGTCTCGTCCTGGGCATTGTGATGGATTGACCCTCCATTTGCAGTACCTTCCCAG
CTGATTAAAGTTCAGCAGTGGTATTGAGGTTTTTCGAATATTTATATAGAAAAAAAGTCTTTTCACATGACAAAT
GACACTCTCACACCAGTCTTAGCCCTAGTAGTTTTTTAGGTTGGACCAGAGGAAGCAGGTTAAATGAGACCTGTC
CTCTGCTGCACTCAGAAAAAATAGGCAGTCCCTGATGCTCAGATCTTAGCCTTGATATTAATAGTTGAGACCACC
TACCCACAATGCAGCCTATACTCCCAAGACTACAAAGTTACCATCGCAAAGGAAAGGTTATTCCAGTAAAAGGAA
ATAGTTTTCTCAACCATTTAAAAATATTCTTCTGAACTCATCAAAGTAGAAGAGCCCCCAACCTTTTCTCTCTGC
CTTCAAGAAGGCAGACATTTGGTATGATTTAGCATCAACAACACATTTATGAGTATATGTAAGTAATCAGAGGGG
CAAATGCCACTTGTTATTCCTCCCAAGTTTTCCAAGCAAGTACACACAGATCTCTGGTAGGATTAGGGGCCACTT
GTGTTTCCGGCTTATTTTAGTCGACTTGTCAGCAAGTTTGATGCCTAGTCTATCTGACATGGCCCAGTAGAACAG
GGCATTGATGGATCACATGAGATGGTAGAAGGAACATCATCACATACCCTCTCACAGAGAAAATTATCAAAGAA
CCAGAAATTATATCTGTTTTGGAGCAAGAGTGTCATAATGTTTCAGGGTAGTCAAAATAAACATAAATTATCTCC
TCTAGATGAGTGGCGATGTTGGCTGATTTGGGTCTGCCATTGACAGAATGTCAAATAAAAAGGAATTAGCTAGAA
TATGACCATTAAATGTGCTTCTGAAATATATTTTGAGATAGGTTTAGAATGTCA
```

FIGURE 450

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPHN

LSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNHICSVQGDAFQKLRRVKELTLSSNQIT

QLPNTTFRPMPNLRSVDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRSLKF

LDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLISLHSLCLRRNKVAIVVSSL

DWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLTSITLAGNL

WDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDAVYAFHLCEDGAEPTSGHLLSAV

TNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMALIFSFLI

VVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNHIEGALVII

NEYGSCTCHQQPARECEV

FIGURE 451

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCTCCAGTCACCCTCCCGCCGTTAC
CCGCGGCGCGCCCGAGGGAGTCTCCTCCAGACCCTCCCTCCCGTTGCTCCAAACTAATACGGA
CTGAACGGATCGCTGCGAGGGTGGGAGAGAAAATTAGGGGGAGAAAGGACAGAGAGAGCAACT
ACCATCCATAGCCAGATAGATTATCTTACACTGAACTGATCAAGTACTTTGAAA**ATG**ACTTCG
AAATTTATCTTGGTGTCCTTCATACTTGCTGCACTGAGTCTTTCAACCACCTTTTCTCTCCAA
CTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACTTATATAAA
GTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAAT
GTTTTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAACTGGCCTCTTTGCAGAGAAT
CATGGGATTGTTGCAAATGATATGTTTGATCCTATTCGGAACAAATCTTTCTCCTTGGATCAC
ATGAATATTTATGATTCCAAGTTTTGGGAAGAAGCGACACCAATATGGATCACAAACCAGAGG
GCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGAACAGATGTAAAAATACATAAGCGCTTT
CCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAGAGTTGCCAAAATTGTT
GAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTGATGAC
ATGGGCCACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTTCAGATATTGACAAG
AAGTTAGGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATC
ATCACAAGTGATCATGGAATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTAC
CTGGATAAAGACCACTATACCCTGATTGATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAA
GGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCTCATCCTAATCTTACTGTTTACAAA
AAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCAACCAATCATAGCA
GTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACCAC
GGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTTTAGCCCATGGTCCTGCCTTCAGA
AAGAATTTCTCAAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTC
AATATCACTGCCATGCCACACAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCA
ATGCCAAGGGTGGTCCCTTATACACAGAGTACTATACTCCTCCCTGGTAGTGTTAAACCAGCA
GAATATGACCAAGAGGGGTCATACCCTTATTTCATAGGGGTCTCTCTTGGCAGCATTATAGTG
ATTGTATTTTTTGTAATTTTCATTAAGCATTTAATTCACAGTCAAATACCTGCCTTACAAGAT
ATGCATGCTGAAATAGCTCAACCATTATTACAAGCC**TAA**TGTTACTTTGAAGTGGATTTGCAT
ATTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTCTGGGAAACCAGTT
CCAAACATCTGCAGAAACCATTAAGCAGTTACATATTTAGGTATACACACACACACACACACA
CACATACACACACGGACCAAAATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGT
CTCCATTGTTCACTGTAGCATAGGGATAGATAAGATCCTGCTTTATTTGGACTTGGCGCAGAT
AATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATATATTGCACTTTAAATTTCT
CTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAACTTGATTG
AAAATGACAACTTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACTGAAGGA
AATTTCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAA
GGTGATAAGTGTTGAAAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCC
CAACAGCAGAATGCAACTGTGGGCATTTCTTGTCTTATTTCTTTCCAGAGAACGTGGTTTTCA
TTTATTTTTCCCTCAAAAGAGAGTCAAATACTGACAGATTCGTTCTAAATATATTGTTTCTGT
CATAAAATTATTGTGATTTCCTGATGAGTCATATTACTGTGATTTTCATAATAATGAAGACAC
CATGAATATACTTTTCTTCTATATAGTTCAGCAATGGCCTGAATAGAAGCAACCAGGCACCAT
CTCAGCAATGTTTTCTCTTGTTTGTAATTATTTGCTCCTTTGAAAATTAAATCACTATTAATT
ACATTAAAAATCAAATTGGATAAAAAAAAAAAAAAAAAA

FIGURE 452

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVKQ

VTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIWIT

NQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLYWED

PDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEERLIEL

DQYLDKDHYTLIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYNSRIQP

IIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTDLYPLLC

HLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYFIGVSLGS

IIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Important features:

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 453

GGCCGCCTGGAATTGTGGGAGTTGTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGAC
TATGGCTCCCCAGAGCCTGCCTTCATCTAGGATGGCTCCTCTGGGCATGCTGCTTGGGCTGCT
GATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGCCCTGACCAA
CCCAGAGAAGAGCAGCACCAAAGAAACGGAGAGAAAAGAAACCAAAGCCGAGGAGGAGCTGGA
TGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGGCAGGC
TGTCCCTGCAGGATCCCACGTACGGCTGAATCTTCAGACTGGGGAAAGAGAGGCAAAACTCCA
ATATGAGGACAAGTTCCGAAATAATTTGAAAGGCAAAAGGCTGGATATCAACACCAACACCTA
CACATCTCAGGATCTCAAGAGTGCACTGGCAAAATTCAAGGAGGGGGCAGAGATGGAGAGTTC
AAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGAGGAACTGAA
GAAAGACTTTGATGAGCTGAATGTTGTCATTGAGACTGACATGCAGATCATGGTACGGCTGAT
CAACAAGTTCAATAGTTCCAGCTCCAGTTTGGAAGAGAAGATTGCTGCGCTCTTTGATCTTGA
ATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTCCTTTGGTGGTCTTCAAGTGGT
GATCAATGGGCTGAACAGCACAGAGCCCCTCGTGAAGGAGTATGCTGCGTTTGTGCTGGGCGC
TGCCTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGAGCCCTGCAGAAGCT
GCTGGTCATCCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAAGGTCCTGTTTGCACTGTG
CTCCCTGCTGCGCCACTTCCCCTATGCCCAGCGGCAGTTCCTGAAGCTCGGGGGGCTGCAGGT
CCTGAGGACCCTGGTGCAGGAGAAGGGCACGGAGGTGCTCGCCGTGCGCGTGGTCACACTGCT
CTACGACCTGGTCACGGAGAAGATGTTCGCCGAGGAGGAGGCTGAGCTGACCCAGGAGATGTC
CCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAGGCCTGTGGGAACAGGGCTG
GTGCGAGATCACGGCCCACCTCCTGGCGCTGCCCGAGCATGATGCCCGTGAGAAGGTGCTGCA
GACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCGTCAGGACCCCCAGCTCGGCAG
GACACTGGCCAGCCTGCAGGCTGAGTACCAGGTGCTGGCCAGCCTGGAGCTGCAGGATGGTGA
GGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAACAGCTTGCTGAAGGAGCTGAGA**TG**
AGGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTGAGGCTGAGGGGTGCCAGCGTGGGTG
GGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCTTGGCCATTAAATGGAAACCTGAAGG
CCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 454

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEELD
AEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTNTY
TSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMVRLI
NKFNSSSSSLEEKIAALFDLEYYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAFVLGA
AFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKLGGLQV
LRTLVQEKGTEVLAVRVVTLLYDLVTEKMFAEEEAELTQEMSPEKLQQYRQVHLLPGLWEQGW
CEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLASLELQDGE
DEGYFQELLGSVNSLLKELR

Important features:

Signal peptide:

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 455

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCCAGAGCCCAGGAGGAGGCAGT

GGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCCACCCCCTA

CCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGGTGT

CTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCTAGAC

TCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGTCCTTG

TGGTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCC**ATG**GCTACAGCAAGACCCCCCTGGA

TGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTCACAGAGCATGTTCTCGCCA

ACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCAGGACCTGG

GAGCTGGGGCCGGGGAAGACGCCCGGTCGGATGACAGCAGCAGCCGCATCATCAATGGATCCG

ACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAACCAGCTCTACT

GCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAGGAAGAAAGTTT

TCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGCAGCAGATGTTCC

AGGGGGTCAAATCCATCCCCCACCCTGGCTACTCCCACCCTGGCCACTCTAACGACCTCATGC

TCATCAAACTGAACAGAAGAATTCGTCCCACTAAAGATGTCAGACCCATCAACGTCTCCTCTC

ATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAACCAAGAGCCCCCAAG

TGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGTCAGAAAAGGTGCGAGG

ATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAGAGACT

CCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCCT

GGGGAGATTACCCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTCA

CCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCATCCCAGGACTCAGCACACCGG

CATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAGACCCTCATTCCTTCCCAGAGA

TGTTGAGAATGTTCATCTCTCCAGCCCCTGACCCCATGTCTCCTGGACTCAGGGTCTGCTTCC

CCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGGGAACAATTTCCAAAACTGTCCAG

GGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGCTCAGGGCCCATCCCTT

CTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAACTGAGAAGTGGAAAAAAAA

FIGURE 456

MATARPPWMWVLCALITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDDS
SSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLSPV
YESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCPSAGTKCLVSG
WGTTKSPQVHFPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKAGRDSCQGDSGGPVVCN
GSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

FIGURE 457

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCTTGCTTCCTGAACT
AGCTCACAGTAGCCCGGCGGCCCAGGGCAATCCGACCACATTTCACTCTCACCGCTGTAGGAATCCAG**ATG**CAGG
CCAAGTACAGCAGCACGAGGGACATGCTGGATGATGATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTG
CCACAACTCGGCATCCAGAGCCCCGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGA
CCCTGCTGACTTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACCAGC
TCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCCCAAGAGTTGCAATCTC
TTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGCTGAAAAACTCTGTCGTGAGCTGTATAACA
AAGCTGGAGCACACAGGTGCAGCCCTTGTACAGAACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATA
AAGACAGCAAAAGTTGGGAGGACTGTAAATATTTCTGCCTTAGTGAAAACTCTACCATGCTGAAGATAAACAAAC
AAGAAGACCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTTTGCGCC
CTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACTTCTGAACTGTTCCATATTATAATAGATG
TCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGGGATGATCTTCTCAAAGGACTGCAAAGAATTGA
AGCGTTGTGTCTGTGAGAGAAGGGCAGGAATGGTGAAGCCAGAGAGCCTCCATGTCCCCCCTGAAACATTAGGCG
AAGGTGAC**TGA**TTCGCCCTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAG
ACATTGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTGTTTCCTGTT
CAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACAAGAAGTCTTATTTACATGCCAC
CAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTGGCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAA
TGTCTAATATCACCTCCCTGTTTTCATGTCTTCCTTACACTTGGTGGAATAAGAAACTTTTTGAAGTAGAGGAAA
TACATTGAGGTAACATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTAC
CAGCAAATACACAAGGAATTCTTTTTGTTTGTTTCAGTTCATACTAGTCCCTTCCCAATCCATCAGTAAAGACCC
CATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAGAATCTCAAATCTCAATGCCTTATAA
GCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTGTCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATAT
ATCCCCATCTCCGTTTCATATCAGAACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGT
GAGCCTCTTCATCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAGGT
ACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

FIGURE 458

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLLI
GLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRELY
NKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAASQSY
SEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKDCKEL
KRCVCERRAGMVKPESLHVPPETLGEGD

FIGURE 459

GTTGATGGCAAACTTCCTCAAAGGAGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCCAC

TGGCGGCCCGCAACACTCCGTCTCACCCTCTGGGCCCACTGCATCTAGAGGAGGGCCGTCTGT

GAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGTGGT

CAGCTGGGTCAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAGGGAA

GTGGGAGCCTCGAGCCCTCGGGTGGAAGCTGACCCCAAGCCACCCTTCACCTGGACAGG**ATG**A

GAGTGTCAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGATGTTTA

TTCGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCAGCCTCGC

CCACCAAGGAGATCCAGGTTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCCCAGCCAACT

ACTTTGCGTTTAAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGTGCTTTGAAG

ACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATCGCCCTGGTGA

ATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGGAGATGTTATGC

ACCTAGTGAAATTCCTTAAAGAAATTCCGGGGGGTGCACTGGTGCTGGTGGCCTCCTACGACG

ATCCAGGGACCAAAATGAACGATGAAAGCAGGAAACTCTTCTCTGACTTGGGGAGTTCCTACG

CAAAACAACTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGACCTCAGGGGTAAAA

GCCCCTTTGAGCAGTTCTTAAAGAACAGCCCAGACACAAACAAATACGAGGGATGGCCAGAGC

TGCTGGAGATGGAGGGCTGCATGCCCCCGAAGCCATTT**TAG**GGTGGCTGTGGCTCTTCCTCAG

CCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCGGCAGGGGCTGAGGAGGA

GGAGCAGGGGGTGCTGCGTGGAAGGTGCTGCAGGTCCTTGCACGCTGTGTCGCGCCTCTCCTC

CTCGGAAACAGAACCCTCCCACAGCACATCCTACCCGGAAGACCAGCCTCAGAGGGTCCTTCT

GGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTCGTCAGGGACTGCTGACGGCTGGTCCTG

AGGAAGGACAAACTGCCCAGACTTGAGCCCAATTAAATTTTATTTTTGCTGGTTTTGAAAAAA

AAAAAAAAAAAAAA

FIGURE 460

MRVSGVLRLLALIFAIVTTWMFIRSYMSFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCPA
NYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSGDV
MHLVKFLKEIPGGALVLVASYDDPGTKMNDESRKLFSDLGSSYAKQLGFRDSWVFIGAKDLRG
KSPFEQFLKNSPDTNKYEGWPELLEMEGCMPPKPF

Important features:

Signal peptide:

amino acids 1-15

ATP/GTP-binding site motif A (P-loop).

amino acids 184-191

N-glycosylation site.

amino acids 107-110

FIGURE 461

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGGA
AACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCCTC
TAGAACCCGACCCACCACC**ATG**AGGTCCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGGCGT
CCAGTGGTCCTTGCTTCTGGCTGTCCTGGTCTTCTTTCTCTTCGCCTTGCCCTCTTTTATTAA
GGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCTCTACA
GTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTATGCAGA
GCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCCACACCACCGGAGA
CAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCACACAGCACA
GAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAACCATGGTGAACACACTGTCACCCAGAGG
GCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCCAGGACACAAA
GACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACGGTGTCAGAGAA
GCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAGTCAGCACAGAATGCT
GGCTCCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCACAGCAGTCATCCC
ACCTAAGGAGAAGAAACCTCAGGCCACCCCACCCCCTGCCCCTTTCCAGAGCCCCACGACGCA
GAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTGGGATTTTGAGGAAAA
ATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTGTGAAGATCAAAGCCTC
CAAGTCGCTGTGGCTCCAGAAACTCTTTCTGCCCAACCTCACTCTCTTCCTGGACTCCAGACA
CTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACCCTTTGGCTTCATGGAGCT
CAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCTCCAGTGCCCCAGCAGCAGCTGCT
CCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGTGCCGTGGTGGGCAACGGGGG
CATCCTGAACAACTCCCACATGGGCCAGGAGATAGACAGTCACGACTACGTGTTCCGATTGAG
CGGAGCTCTCATTAAAGGCTACAACAGGATGTGGGGACTCGGACATCCTTCTACGGCTTTAC
CGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAATCGGGGTTTCAAGAACGTGCCTCT
TGGGAAGGACGTCCGCTACTTGCACTTCCTGGAAGGCACCCGGGACTATGAGTGGCTGGAAGC
ACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTTTCTGGTTCAGGCACAGACCCCAGGA
AGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTGTTGCTGCACCCAGACTTTCTCCGATA
CATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGATGGTGCCCACTGGAGGATATACCGCCC
CACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTCAGCTCTGTGACCAGGTGAGTGCTTATGG
CTTCATCACTGAGGGCCATGAGCGCTTTTCTGATCACTACTATGATACATCATGGAAGCGGCT
GATCTTTTACATAAACCATGACTTCAAGCTGGAGAGAGAAGTCTGGAAGCGGCTACACGATGA
AGGGATAATCCGGCTGTACCAGCGTCCTGGTCCCGGAACTGCCAAAGCCAAGAAC**TGA**CCGGG
GCCAGGGCTGCCATGGTCTCCTTGCCTGCTCCAAGGCACAGGATACAGTGGGAATCTTGAGAC
TCTTTGGCCATTTCCCATGGCTCAGACTAAGCTCCAAGCCCTTCAGGAGTTCCAAGGGAACAC
TTGAACCATGGACAAGACTCTCTCAAGATGGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCA
GTACATTGCTGTAGGTCCTGAGGCCAGGGATTTTTAATTAAATGGGGTGATGGGTGGCCAATA
CCACAATTCCTGCTGAAAAACACTCTTCCAGTCCAAAAGCTTCTTGATACAGAAAAAGAGCC
TGGATTTACAGAACATATAGATCTGGTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCT
TGAAGGTATTACTTAACTTCACTACAGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTC
TAGAAGGGTCTATACTTGTCCTTGTCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAAC
TGATAATAATACAAATGATTGTTGTCCATGGAAAGGCAAATAAATTTTCTACAGTGAAAAAAA
AAAAAAAA

FIGURE 462

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKPK

SQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAWKS

PEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQGKAA

TTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRNQRLK

AANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHFNQSEW

DRLEHFAPPFGFMELNYSLVQKVVTRFPPVPQQQLLLASLPAGSLRCITCAVVGNGGILNNSH

MGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVPLGKDVRY

LHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFLRYMKNRFL

RSKTLDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSWKRLIFYINH

DFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Important features:

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

FIGURE 463

```
GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGCGGCGCAAGGGTGAGGGCGGCCCCAGAAC
CCCAGGTAGGTAGAGCAAGAAGATGGTGTTTCTGCCCCTCAAATGGTCCCTTGCAACCATGTC
ATTTCTACTTTCCTCACTGTTGGCTCTCTTAACTGTGTCCACTCCTTCATGGTGTCAGAGCAC
TGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCCTTGGAATAAAATACGACTTCCTGA
GTACGTCATCCCAGTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACCTTCTG
GGGAACCACGAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCCTGCATAGTCA
CCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGGAAGAACC
CCTGCAGGTCCTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAGCCCCTCCT
TGTCGGGCTCCCGTACACAGTTGTCATTCACTATGCTGGCAATCTTTCGGAGACTTTCCACGG
ATTTTACAAAAGCACCTACAGAACCAAGGAAGGGGAACTGAGGATACTAGCATCAACACAATT
TGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGATGAACCTGCCTTCAAAGCAAGTTT
CTCAATCAAAATTAGAAGAGAGCCAAGGCACCTAGCCATCTCCAATATGCCATTGGTGAAATC
TGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGAAGATGAGCACCTA
TCTGGTGGCCTTCATCATTTCAGATTTTGAGTCTGTCAGCAAGATAACCAAGAGTGGAGTCAA
GGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGCACTGGATGCTGCGGT
GACTCTTCTAGAATTTTATGAGGATTATTTCAGCATACCGTATCCCCTACCCAAACAAGATCT
TGCTGCTATTCCCGACTTTCAGTCTGGTGCTATGGAAAACTGGGGACTGACAACATATAGAGA
ATCTGCTCTGTTGTTTGATGCAGAAAAGTCTTCTGCATCAAGTAAGCTTGGCATCACAGTGAC
TGTGGCCCATGAACTGGCCCACCAGTGGTTTGGGAACCTGGTCACTATGGAATGGTGGAATGA
TCTTTGGCTAAATGAAGGATTTGCCAAATTTATGGAGTTTGTGTCTGTCAGTGTGACCCATCC
TGAACTGAAAGTTGGAGATTATTTCTTTGGCAAATGTTTTGACGCAATGGAGGTAGATGCTTT
AAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTGCTCAGATCCGGGAGATGTTTGA
TGATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATGCTAAGGGAGTATCTTAGCGCTGA
CGCATTTAAAAGTGGTATTGTACAGTATCTCCAGAAGCATAGCTATAAAAAATACAAAAAACGA
GGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAGATGGTGTAAAAGGGATGGATGGCTT
TTGCTCTAGAAGTCAACATTCATCTTCATCCTCACATTGGCATCAGGAAGGGGTGGATGTGAA
AACCATGATGAACACTTGGACACTGCAGAGGGGTTTTCCCCTAATAACCATCACAGTGAGGGG
GAGGAATGTACACATGAAGCAAGAGCACTACATGAAGGGCTCTGACGGCGCCCCGGACACTGG
GTACCTGTGGCATGTTCCATTGACATTCATCACCAGCAAATCCAACATGGTCCATCGATTTTT
GCTAAAAACAAAAACAGATGTGCTCATCCTCCCAGAAGAGGTGGAATGGATCAAATTTAATGT
GGGCATGAATGGCTATTACATTGTGCATTACGAGGATGATGGATGGGACTCTTTGACTGGCCT
TTTAAAAGGAACACACAGCAGTCAGCAGTAATGATCGGGCAAGTCTCATTAACAATGCATT
TCAGCTCGTCAGCATTGGGAAGCTGTCCATTGAAAAGGCCTTGGATTTATCCCTGTACTTGAA
ACATGAAACTGAAATTATGCCCGTGTTTCAAGGTTTGAATGAGCTGATTCCTATGTATAAGTT
AATGGAGAAAAGAGATATGAATGAAGTGGAAACTCAATTCAAGGCCTTCCTCATCAGGCTGCT
AAGGGACCTCATTGATAAGCAGACATGGACAGACGAGGGCTCAGTCTCAGAGCAAATGCTGCG
GAGTGAACTACTACTCCTCGCCTGTGTGCACAACTATCAGCCGTGCGTACAGAGGGCAGAAGG
CTATTTCAGAAAGTGGAAGGAATCCAATGGAAACTTGAGCCTGCCTGTCGACGTGACCTTGGC
AGTGTTTGCTGTGGGGGCCCAGAGCACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTT
TTCTTTGTCCAGTACTGAGAAAAGCCAAATTGAATTTGCCCTCTGCAGAACCCAAAATAAGGA
AAAGCTTCAATGGCTACTAGATGAAAGCTTTAAGGGAGATAAAATAAAAACTCAGGAGTTTCC
ACAAATTCTTACACTCATTGGCAGGAACCCAGTAGGATACCCACTGGCCTGGCAATTTCTGAG
GAAAAACTGGAACAAACTTGTACAAAAGTTTGAACTTGGCTCATCTTCCATAGCCCACATGGT
AATGGGTACAACAAATCAATTCTCCACAAGAACACGGCTTGAAGAGGTAAAAGGATTCTTCAG
CTCTTTGAAAGAAAATGGTTCTCAGCTCCGTTGTGTGTCCAACAGACAATTGAAACCATTGAAGA
AAACATCGGTTGGATGGATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCT
TGAACGTATGTAAAAATTCCTCCCTTGCCCGGTTCCTGTTATCTCTAATCACCAACATTTTGT
TGAGTGTATTTTCAAACTAGAGATGGCTGTTTTGGCTCCAACTGGAGATACTTTTTTTCCCTTC
AACTCATTTTTTGACTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTT
CATGAATGGGCTATCGCTACCATGTGTTTTGTTCATCACAGGTGTTGCCCTGCAACGTAAACC
CAAGTGTTGGGTTCCCTGCCACAGAAGAATAAAGTACCTTATTCTTCTCAAAAAAAAAAAAAA
AAAAAAAAAAAA
```

FIGURE 464

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVHY
DLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLEHP
PQEQIALLAPEPLLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTAARM
AFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVAFIIS
DFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAAIPDFQ
SGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDLWLNEGF
AKFMEFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFDDVSYDKG
ACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDGFCSRSQHS
SSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPDTGYLWHVPL
TFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSLTGLLKGTHTA
VSSNDRASLINNAFQLVSIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIPMYKLMEKRDMN
EVETQFKAFLIRLLRDLIDKQTWTDEGSVSEQMLRSELLLLACVHNYQPCVQRAEGYFRKWKE
SNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALCRTQNKEKLQWLLD
ESFKGDKIKTQEFPQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGSSSIAHMVMGTTNQF
STRTRLEEVKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIRVWLQSEKLERM

Important features:
Signal peptide:
amino acids 1-34

N-glycosylation sites:
amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:
amino acids 350-360

FIGURE 465

CAGCCACAGACGGGTC**ATG**AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCACT

GCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCCGA

CCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGACAC

GTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGGAGGC

CAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATCTCCTA

CACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCTTTGGGC

CCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCCAGTCTGCTTGTCTATGGAAGGCTG

TCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGATGGCCTCCT

CAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCCCCAGCCAGG

TTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCGTGGGTATGACTGAGAACTGCAATAG

GAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAACTTGGCTCAAGA

ACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGTGTGTCAGGAGAC

GCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAGGCTGCAGCACTGT

TGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCTCCTGGGGTGCTTGTGGCCTC

CTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAGCAGCGTTCTGCTGAA

CTCCCTCCCTCCTCAAGCTGCCCCTGTCCCAGGAGACCGGCAGTGTCCTACCTGTGTGCAGCC

CCTTGGAACCTGTTCAAGTGGCTCCCCCCGAATGACCTGCCCCAGGGGCGCCACTCATTGTTA

TGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAATGAGCATTCAGGGCTGCGT

GGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCGGGATCTTCTCTGCGCGTGA

GAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGTGGGGCTGAGGGCCTGGAGTC

TCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTGGTGGGGAGTGGTTTGCCCTTC

CTGC**TAA**CTCTATTACCCCACGATTCTTCACCGCTGCTGACCACCCACACTCAACCTCCCTC

TGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTCCCATTCTGTCCATGAATCATCTT

CCCCACACACAATCATTCATATCTACTCACCTAACAGCAACACTGGGGAGAGCCTGGAGCATC

CGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTGGCTGCATGTATCTGATAATACAGAC

CCTGTCCTTTCA

FIGURE 466

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLIE
SGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQPPA
DPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCNLLN
GTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETLLLID
VGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLNSLPPQ
AAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGCVAQPSS
FLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVCPSC

FIGURE 467

```
GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGAGCACCACCGGAGCCCTTGAGACATCCTTG
AGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTGCAGGATGATGGTGGCCCTTCG
AGGAGCTTCTGCATTGCTGGTTCTGTTCCTTGCAGCTTTTCTGCCCCCGCCGCAGTGTACCCA
GGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTCGAGTCTTGGAGCAAGGGCTGGAAAA
ATGTACCCAAGCAACGAGGGCATACATTCAAGAATTCCAAGAGTTCTCAAAAAATATATCTGT
CATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACTTGGCACT
GAGAGTTGAACGTGCCCAACGGGAGATTGACTACATACAATACCTTCGAGAGGCTGACGAGTG
CATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAAGAGAAAA
GATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTCTTTGAAAAT
AGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATAACTCTCCAAA
GGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAACATACGGGCATT
CATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTCCTGGCAGGGAAC
AGGCCAAGTGATCTACAAAGGTTTTCTATTTTTTCATAACCAAGCAACTTCTAATGAGATAAT
CAAATATAACCTGCAGAAGAGGACTGTGGAAGATCGAATGCTGCTCCCAGGAGGGGTAGGCCG
AGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGTGGATGAGCATGGGCT
CTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCACAAAGATTGAGCCGGG
CACACTGGGAGTGGAGCATTCATGGGATACCCCATGCAGAAGCCAGGATGCTGAAGCCTCATT
CCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCAGGGCCCTCATCGCATCAC
CTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGCCCAACTTGTTCTTCCCCAA
GAGACCAAGAAGTCACTCCATGATCCATTACAACCCCAGAGATAAGCAGCTCTATGCCTGGAA
TGAAGGAAACCAGATCATTTACAAACTCCAGACAAAGAGAAAGCTGCCTCTGAAGTAATGCAT
TACAGCTGTGAGAAGAGCACTGTGGCTTTGGCAGCTGTTCTACAGGACAGTGAGGCTATAGC
CCCTTCACAATATAGTATCCCTCTAATCACACACAGGAAGAGTGTGTAGAAGTGGAAATACGT
ATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTCCAAGAGCTTAGATGAGAGCATATC
ATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAAACCTCCTGGCTCTCAAGGATGACCAC
ATTCTGATACAGCCTACTTCAAGCCTTTTGTTTTACTGCTCCCCAGCATTTACTGTAACTCTG
CCATCTTCCCTCCCACAATTAGAGTTGTATGCCAGCCCCTAATATTCACCACTGGCTTTTCTC
TCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTTCAAATGTCTATTGATATTCTCCCATTTT
CACTGCCCAACTAAAATACTATTAATATTTCTTTCTTTTCTTTTCTTTTTTTTGAGACAAGGT
CTCACTATGTTGCCCAGGCTGGTCTCAAACTCCAGAGCTCAAGAGATCCTCCTGCCTCAGCCT
CCTAAGTACCTGGGATTACAGGCATGTGCCACCACACCTGGCTTAAAATACTATTTCTTATTG
AGGTTTAACCTCTATTTCCCCTAGCCCTGTCCTTCCACTAAGCTTGGTAGATGTAATAATAAA
GTGAAAATATTAACATTTGAATATCGCTTTCCAGGTGTGGAGTGTTTGCACATCATTGAATTC
TCGTTTCACCTTTGTGAAACATGCACAAGTCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGT
AACACAATTACAAAGTGAAAGATACAGCTAGAAAATACTACAAATCCCATAGTTTTTCCATTG
CCCAAGGAAGCATCAAATACGTATGTTTGTTCACCTACTCTTATAGTCAATGCGTTCATCGTT
TCAGCCTAAAAATAATAGTCTGTCCCTTTAGCCAGTTTTCATGTCTGCACAAGACCTTTCAAT
AGGCCTTTCAAATGATAATTCCTCCAGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCC
TCCTCTTGTCTTGCTGTCCTCTGTTTCTCTCTTTCTGCTTTAAATTCAATAAAAGTGACACTG
AGCAAAAAAAAAAAAAAAA
```

FIGURE 468

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFRVLEQGLEKCTQATRAYIQEFQEF
SKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLLQE
AEEEKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNSPKVYLLIGSRNNTVWEF
ANIRAFMEDNTKPAPRKQILTLSWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVEDRMLL
PGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKIEPGTLGVEHSWDTPCRSQ
DAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIHYNPRDK
QLYAWNEGNQIIYKLQTKRKLPLK

FIGURE 469

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGGC

AGTGTTTTGCCTTCACCCCAAGTGACC**ATG**AGAGGTGCCACGCGAGTCTCAATCATGCTCCTC

CTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGTGGG

GCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCTGGGG

CGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCAAGCAC

CACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTACCGCTGC

TCCATGGACTTGAAGAACATCAATTTT**TAG**GCGCTTGCCTGGTCTCAGGATACCCACCATCCT

TTTCCTGAGCACAGCCTGGATTTTTATTTCTGCCATGAAACCCAGCTCCCATGACTCTCCCAG

TCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAGGCAGACATA

CCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGGCTGTGGTGTG

AAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTAAATGGCAGAAA

GGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTGCCCCTCTCCCCA

CATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTGGGTGCATTGCTCAGAG

TCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGGACCAATTTGTGGGT

AGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGACTCAAGATTGGCTCTT

CCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCAGGGAGGCCAATCAGCC

CCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGACCTGTGACCTTCTGCCA

GAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGTTAACCACTGAAGCCCCCA

ATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAATCTAATCTGATATTGACAT

ATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACTCCTTTCCAAGGATCAGCCCTGAGAGCAG

GTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGGTGGGAGCAAGGGACAGGGAGC

AGGGCAGGGGCTGAAAGGGGCACTGATTCAGACCAGGGAGGCAACTACACACCAACATGCTGG

CTTTAGAATAAAAGCACCAACTGAAAAAA

FIGURE 470

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHPG
SHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Important feratures:

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

FIGURE 471

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAA**ATG**TCTTTCC

TCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGCTG

CTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGGCCC

TGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAAGCAA

AGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTTCCTCT

GTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCGTCCCCC

TCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTATTTCAAAG

GAGAAATCTTCCTGGATGAAAGAAAAAGTTCTATGGTCCACAAAGGCGGAAGATGATGTTTA

TGGGATTTATCCGTCTGGGAGTGTGGTACAACTTCTTCCGAGCCTGGAACGGAGGCTTCTCTG

GAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCGTGGTGGGATCAGGAAAGCAGG

GCATTCTTCTTGAGCACCGAGAAAAGAATTTGGAGACAAAGTAAACCTACTTTCTGTTCTGG

AAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAA**TGA**TTGTGTGAAACTG

CCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTCATGGGATGTATTGTTTCCACTCGTG

TCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTATTAATGTATTTTAA

TATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGACAAAAATCTGAAAAA

CTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAATTGACTGCCAGGCTGG

GTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGAGCAAGTCACTTGAG

GTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAA

ATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCGGGAGGCTGAGGCAGGAGAA

TCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCACACCACTGTATTCCAGCCTG

GGTGACTGAGACTCTAACTAA

FIGURE 472

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKAALEYLEDIDLKTLEKEPRT
FKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDFQP
YFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEGFILGGVFVVGS
GKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 473

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTTT

TGGGATTTTAATTTTCAAACACAGCAGA**ATG**ACATTTTTTCTGTCACTATTATTATTGTTGGT

ATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCTATC

AAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTCCTAC

CAAAGCTGTCAAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGAGGGTT

AATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACT**TAG**ATTCAATGATTGTA

AATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATATATACAA

TAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACTTTATTAAT

TTTTAAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGATCATATAAT

TTGATACAAATAAAAGAAAGTGTTCTCTCCCCTTACAGAATTGACATTTTAAATGCGATACA

GTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAAAGAAGGGAAAA

TGTTGCCAAGGAAAAAAAAA

FIGURE 474

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGKG
IVKGRNLDSRGLILGAEAWGRGVKKNT

FIGURE 475

GACAGTGGAGGGCAGTGGAGAGGACCGCGCTGTCCTGCTGTCACCAAGAGCTGGAGACACCAT

CTCCCACCGAGAGTC**ATG**GCCCCATTGGCCCTGCACCTCCTCGTCCTCGTCCCCATCCTCCTC

AGCCTGGTGGCCTCCCAGGACTGGAAGGCTGAACGCAGCCAAGACCCCTTCGAGAAATGCATG

CAGGATCCTGACTATGAGCAGCTGCTCAAGGTGGTGACCTGGGGGCTCAATCGGACCCTGAAG

CCCCAGAGGGTGATTGTGGTTGGCGCTGGTGTGGCCGGGCTGGTGGCCGCCAAGGTGCTCAGC

GATGCTGGACACAAGGTCACCATCCTGGAGGCAGATAACAGGATCGGGGGCCGCATCTTCACC

TACCGGGACCAGAACACGGGCTGGATTGGGGAGCTGGGAGCCATGCGCATGCCCAGCTCTCAC

AGGATCCTCCACAAGCTCTGCCAGGGCCTGGGGCTCAACCTGACCAAGTTCACCCAGTACGAC

AAGAACACGTGGACGGAGGTGCACGAAGTGAAGCTGCGCAACTATGTGGTGGAGAAGGTGCCC

GAGAAGCTGGGCTACGCCTTGCGTCCCCAGGAAAAGGGCCACTCGCCCGAAGACATCTACCAG

ATGGCTCTCAACCAGGCCCTCAAAGACCTCAAGGCACTGGGCTGCAGAAAGGCGATGAAGAAG

TTTGAAAGGCACACGCTCTTGGAATATCTTCTCGGGGAGGGGAACCTGAGCCGGCCGGCCGTG

CAGCTTCTGGGAGACGTGATGTCCGAGGATGGCTTCTTCTATCTCAGCTTCGCCGAGGCCCTC

CGGGCCCACAGCTGCCTCAGCGACAGACTCCAGTACAGCCGCATCGTGGGTGGCTGGGACCTG

CTGCCGCGCGCGCTGCTGAGCTCGCTGTCCGGGCTTGTGCTGTTGAACGCGCCCGTGGTGGCG

ATGACCCAGGGACCGCACGATGTGCACGTGCAGATCGAGACCTCTCCCCCGGCGCGGAATCTG

AAGGTGCTGAAGGCCGACGTGGTGCTGCTGACGGCGAGCGGACCGGCGGTGAAGCGCATCACC

TTCTCGCCGCCGCTGCCCCGCCACATGCAGGAGGCGCTGCGGAGGCTGCACTACGTGCCGGCC

ACCAAGGTGTTCCTAAGCTTCCGCAGGCCCTTCTGGCGCGAGGAGCACATTGAAGGCGGCCAC

TCAAACACCGATCGCCCGTCGCGCATGATTTTCTACCCGCCGCCGCGCGAGGGCGCGCTGCTG

CTGGCCTCGTACACGTGGTCGGACGCGGCGGCAGCGTTCGCCGGCTTGAGCCGGGAAGAGGCG

TTGCGCTTGGCGCTCGACGACGTGGCGGCATTGCACGGGCCTGTCGTGCGCCAGCTCTGGGAC

GGCACCGGCGTCGTCAAGCGTTGGGCGGAGGACCAGCACAGCCAGGGTGGCTTTGTGGTACAG

CCGCCGGCGCTCTGGCAAACCGAAAAGGATGACTGGACGGTCCCTTATGGCCGCATCTACTTT

GCCGGCGAGCACACCGCCTACCCGCACGGCTGGGTGGAGACGGCGGTCAAGTCGGCGCTGCGC

GCCGCCATCAAGATCAACAGCCGGAAGGGGCCTGCATCGGACACGGCCAGCCCCGAGGGGCAC

GCATCTGACATGGAGGGGCAGGGGCATGTGCATGGGGTGGCCAGCAGCCCCTCGCATGACCTG

GCAAAGGAAGAAGGCAGCCACCCTCCAGTCCAAGGCCAGTTATCTCTCCAAAACACGACCCAC

ACGAGGACCTCGCAT**TAA**AGTATTTTCGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAA

476/550

FIGURE 476

MAPLALHLLVLVPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVVTWGLNRTLKPQRVI
VVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRILHK
LCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGYALRPQEKGHSPEDIYQMALNQ
ALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSEDGFFYLSFAEALRAHSC
LSDRLQYSRIVGGWDLLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNLKVLKA
DVVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGGHSNTDR
PSRMIFYPPPREGALLLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQLWDGTGVV
KRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKSALRAAIKI
NSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQNTTHTRTSH

Important features:
Signal peptide:
amino acids 1-21

FIGURE 477

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTGCCTGC**AT**
GGACGCTCTGAAGCCACCCTGTCTCTGGAGGAACCACGAGCGAGGGAAGAAGGACAGGGACTCGTGTGGCAGGAA
GAACTCAGAGCCGGGAAGCCCCCATTCACTAGAAGCACTGAGAGATGCGGCCCCCTCGCAGGGTCTGAATTTCCT
GCTGCTGTTCACAAAGATGCTTTTTATCTTTAACTTTTTGTTTTCCCCACTTCCGACCCCGGCGTTGATCTGCAT
CCTGACATTTGGAGCTGCCATCTTCTTGTGGCTGATCACCAGACCTCAACCCGTCTTACCTCTTCTTGACCTGAA
CAATCAGTCTGTGGGAATTGAGGGAGGAGCACGGAAGGGGGTTTCCCAGAAGAACAATGACCTAACAAGTTGCTG
CTTCTCAGATGCCAAGACTATGTATGAGGTTTTCCAAAGAGGACTCGCTGTGTCTGACAATGGGCCCTGCTTGGG
ATATAGAAAACCAAACCAGCCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTC
CTGTCTCTTGCATAAAGGTTATAAATCATCACCAGACCAGTTTGTCGGCATCTTTGCTCAGAATAGGCCAGAGTG
GATCATCTCCGAATTGGCTTGTTACACGTACTCTATGGTAGCTGTACCTCTGTATGACACCTTGGGACCAGAAGC
CATCGTACATATTGTCAACAAGGCTGATATCGCCATGGTGATCTGTGACACACCCCAAAAGGCATTGGTGCTGAT
AGGGAATGTAGAGAAAGGCTTCACCCCGAGCCTGAAGGTGATCATCCTTATGGACCCCTTTGATGATGACCTGAA
GCAAAGAGGGGAGAAGAGTGGAATTGAGATCTTATCCCTATATGATGCTGAGAACCTAGGCAAAGAGCACTTCAG
AAAACCTGTGCCTCCTAGCCCAGAAGACCTGAGCGTCATCTGCTTCACCAGTGGGACCACAGGTGACCCCAAAGG
AGCCATGATAACCCATCAAAATATTGTTTCAAATGCTGCTGCCTTTCTCAAATGTGTGGAGCATGCTTATGAGCC
CACTCCTGATGATGTGGCCATATCCTACCTCCCTCTGGCTCATATGTTTGAGAGGATTGTACAGGCTGTTGTGTA
CAGCTGTGGAGCCAGAGTTGGATTCTTCCAAGGGGATATTCGGTTGCTGGCTGACGACATGAAGACTTTGAAGCC
CACATTGTTTCCCGCGGTGCCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCTT
GAAGAAGTTCTTGTTGAAGCTGGCTGTTTCCAGTAAATTCAAAGACTTCAAAAGGGTATCATCAGGCATGATAG
TTTCTGGGACAAGCTCATCTTTGCAAAGATCCAGGACAGCCTGGGCGGAAGGGTTCGTGTAATTGTCACTGGAGC
TGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTCAGGTGTATGAAGCTTATGGTCA
AACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTCACGTTGGGGTGCCCCTGGC
TTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAACTACTTTACAGTGAATAATGAAGGAGAGGTCTGCAT
CAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAGAAGACACAGGAAGCCCTGGACAGTGATGGCTG
GCTTCACACAGGAGACATTGGTCGCTGGCTCCCGAATGGAACTCTGAAGATCATCGACCGTAAAAAGAACATTTT
CAAGCTGGCCCAAGGAGAATACATTGCACCAGAGAAGATAGAAAATATCTACAACAGGAGTCAACCAGTGTTACA
AATTTTTGTACACGGGGAGAGCTTACGGTCATCCTTAGTAGGAGTGGTGGTTCCTGACACAGATGTACTTCCCTC
ATTTGCAGCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGAACTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTT
AGAAGACTTGCAGAAAATTGGGAAAGAAAGTGGCCTTAAAACTTTTGAACAGGTCAAAGCCATTTTTCTTCATCC
AGAGCCATTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAAATACTT
TCGGACCCAAATTGACAGCCTGTATGAGCACATCCAGGA**TAG**GATAAGGTACTTAAGTACCTGCCGGCCCACTG
TGCACTGCTTGTGAGAAAATGGATTAAAAACTATTCTTACATTTGTTTTGCCTTTCCTCCTATTTTTTTTTAACC
TGTTAAACTCTAAAGCCATAGCTTTTGTTTTATATTGAGACATATAATGTGTAAACTTAGTTCCCAAATAAATCA
ATCCTGTCTTTCCCATCTTCGATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAA
GATCCAGTTTATGTTCTGTGTCCTTCCTCATGATTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGT
CAAAGGGACCCTCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAACTTGCCAACAGTCTCTATGCTTATTTACA
TCTTCTACTGTTCAAACTAAGACATTTTTAAATTCTGAAAAACTGCTTACAATTCATGTTTTCTAGCCACTCCAC
AAACCACTAAAATTTTAGTTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTT
CTGCGTAAATTAAATTGTGTACTGAAGGGAAAAGTTTGATCATACCAAACATTTCCTAAACTCTCTAGTTAGATA
TCTGACTTGGGAGTATTAAAAATTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTTCTTAAAGCATTATTTA
CAGTAGGAACTGGGGAGTAAATCTGTTCCCTACAGTTTGCTGCTGAGCTGGAAGCTGTGGGGGAAGGAGTTGACA
GGTGGGCCCAGTGAACTTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACCTCCTGAACTGGGAACAAAGATCT
ACAGGCAAGCAAGATGCCCACACAACAGGCTTATTTTCTGTGAAGGAACCAACTGATCTCCCCCACCCTTGGATT
AGAGTTCCTGCTCTACCTTACCCACAGATAACACATGTTGTTTCTACTTGTAAATGTAAAGTCTTTAAAATAAAC
TATTACAGATAAAAAA

FIGURE 478

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFLLLFTKMLFIFNF

LFSPLPTPALICILTFGAAIFLWLITRPQPVLPLLDLNNQSVGIEGGARKGVSQKNNDLTSCC

FSDAKTMYEVFQRGLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSSPDQ

FVGIFAQNRPEWIISELACYTYSMVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKALVLIG

NVEKGFTPSLKVIILMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPPSPEDLSVIC

FTSGTTGDPKGAMITHQNIVSNAAAFLKCVEHAYEPTPDDVAISYLPLAHMFERIVQAVVYSC

GARVGFFQGDIRLLADDMKTLKPTLFPAVPRLLNRIYDKVQNEAKTPLKKFLLKLAVSSKFKE

LQKGIIRHDSFWDKLIFAKIQDSLGGRVRVIVTGAAPMSTSVMTFFRAAMGCQVYEAYGQTEC

TGGCTFTLPGDWTSGHVGVPLACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFKGYLKDPEKT

QEALDSDGWLHTGDIGRWLPNGTLKIIDRKKNIFKLAQGEYIAPEKIENIYNRSQPVLQIFVH

GESLRSSLVGVVVPDTDVLPSFAAKLGVKGSFEELCQNQVVREAILEDLQKIGKESGLKTFEQ

VKAIFLHPEPFSIENGLLTPTLKAKRGELSKYFRTQIDSLYEHIQD

Important features:

Type II transmembrane domain:

amino acids 61-80

Putative AMP-binding domain signature.

amino acids 314-325

N-glycosylation site.

amino acids 102-105, 588-591 and 619-622

FIGURE 479

GGAGGCGGAGGCCGCGGCGAGCCGGGCCGAGCAGTGAGGGCCCTAGCGGGGCCCGAGCGGGGC
CCGGGGCCCCTAAGCCATTCCTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAATCC
GGT**ATG**GACGACTGGAAGCCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGGAGC
TGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGCCCTGCGGAGATTCTGTCAGACA
GGGGCCGTGCTTTTCCTGCTGGTGACTGTCATTGTCAATATCAAGTTGATCCTGGACACTCGG
CGAGCCATCAGTGAAGCCAATGAﾑGACCCAGAGCCAGAGCAAGACTATGATGAGGCCCTAGGC
CGCCTGGAGCCCCCACGGCGCAGAGGCAGTGGTCCCCGGCGGGTCCTGGACGTAGAGGTGTAT
TCAAGTCGCAGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGCTGGAGGATGAGGCCCGG
GAGCAGGGCCGGGGCATCCATGTCATTGTCCTCAACCAGGCCACGGGCCACGTGATGGCAAAA
CGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGCTATTCCTCAACATGGTA
GCGCCCGGCCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACCTCAAGGACACA
GCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGCTGGAGGGACACA
TGGGCCTTCGTGGGACGAAAAGGAGGTCCTGTCTTCGGGGAGAAACATTCTAAGTCACCTGCC
CTCTCTTCCTGGGGGGACCCAGTCCTGCTGAAGACAGATGTGCCATTGAGCTCAGCAGAAGAG
GCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGCTTCTGCAGCAAAGTT
GAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCCACACCCATCGAGTTCAGCCCTGACCCA
CTCCCAGACAACAAGGTCCTCAATGTGCCTGTGGCTGTCATTGCAGGGAACCGACCCAATTAC
CTGTACAGGATGCTGCGCTCTCTGCTTTCAGCCCAGGGGGTGTCTCCTCAGATGATAACAGTT
TTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGGCACTGTTTGGTCTGAGGGGCATC
CAGCATACTCCCATCAGCATCAAGAATGCCCGCGTGTCTCAGCACTACAAGGCCAGCCTCACT
GCCACTTTCAACCTGTTTCCGGAGGCCAAGTTTGCTGTGGTTCTGGAAGAGGACCTGGACATT
GCTGTGGATTTTTTCAGTTTCCTGAGCCAATCCATCCACCTACTGGAGGAGGATGACAGCCTG
TACTGCATCTCTGCCTGGAATGACCAGGGGTATGAACACACGGCTGAGGACCCAGCACTACTG
TACCGTGTGGAGACCATGCCTGGGCTGGGCTGGGTGCTCAGGAGGTCCTTGTACAAGGAGGAG
CTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCTGGGATTGGGACATGTGGATGCGGATGCCT
GAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACGTTTCCCGATCCTACCACTTTGGCATC
GTCGGCCTCAACATGAATGGCTACTTTCACGAGGCCTACTTCAAGAAGCACAAGTTCAACACG
GTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCTGAAGAAAGAAGCTTATGAAGTGGAAGTT
CACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACCACAGCAAGAACCCTTGTGAAGACTCTTTC
CTGCCAGACACAGAGGGCCACACCTACGTGGCCTTTATTCGAATGGAGAAAGATGATGACTTC
ACCACCTGGACCCAGCTTGCCAAGTGCCTCCATATCTGGGACCTGGATGTGCGTGGCAACCAT
CGGGGCCTGTGGAGATTGTTTCGGAAGAAGAACCACTTCCTGGTGGTGGGGGTCCCGGCTTCC
CCCTACTCAGTGAAGAAGCCACCCTCAGTCACCCCAATTTTCCTGGAGCCACCCCCAAAGGAG
GAGGGAGCCCCAGGAGCCCCAGAACAGACA**TGA**GACCTCCTCCAGGACCCTGCGGGGCTGGGT
ACTGTGTACCCCCAGGCTGGCTAGCCCTTCCCTCCATCCTGTAGGATTTTGTAGATGCTGGTA
GGGGCTGGGGCTACCTTGTTTTTAACATGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCC
CTGCTCCAACACCCCGTTCCTGAGTTAAAAGTCTATTTATTTACTTCCTTGTTGGAGAAGGGC
AGGAGAGTACCTGGGAATCATTACGATCCCTAGCAGCTCATCCTGCCCTTTGAATACCCTCAC
TTTCCAGGCCTGGCTCAGAATCTAACCTATTTATTGACTGTCCTGAGGGCCTTGAAAACAGGC
CGAACCTGGAGGGCCTGGATTTCTTTTTGGGCTGGAATGCTGCCCTGAGGGTGGGCTGGCTC
TTACTCAGGAAACTGCTGTGCCCAACCCATGGACAGGCCCAGCTGGGGCCCACATGCTGACAC
AGACTCACTCAGAGACCCTTAGACACTGGACCAGGCCTCCTCTCAGCCTTCTCTTTGTCCAGA
TTTCCAAAGCTGGATAAGTTGGTCATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 480

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQRALRRFCQTGAVLFLLVTVIVNIKLILDTRR

AISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEARE

QGRGIHVIVLNQATGHVMAKRVFDTYSPHEDEAMVLFLNMVAPGRVLICTVKDEGSFHLKDTA

KALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKHSKSPALSSWGDPVLLKTDVPLSSAEEA

ECHWADTELNRRRRRFCSKVEGYGSVCSCKDPTPIEFSPDPLPDNKVLNVPVAVIAGNRPNYL

YRMLRSLLSAQGVSPQMITVFIDGYYEEPMDVVALFGLRGIQHTPISIKNARVSQHYKASLTA

TFNLFPEAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSLYCISAWNDQGYEHTAEDPALLY

RVETMPGLGWVLRRSLYKEELEPKWPTPEKLWDWDMWMRMPEQRRGRECIIPDVSRSYHFGIV

GLNMNGYFHEAYFKKHKFNTVPGVQLRNVDSLKKEAYEVEVHRLLSEAEVLDHSKNPCEDSFL

PDTEGHTYVAFIRMEKDDDFTTWTQLAKCLHIWDLDVRGNHRGLWRLFRKKNHFLVVGVPASP

YSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

Important features:

Transmembrane domain:

amino acids 38-55

Homologous region to Mouse GNT1

amino acids 229-660

FIGURE 481

GAAAGA**ATG**TTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACCA

GGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCATAT

GCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAAGTT

CCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGTATCA

TTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGCAATCA

GCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAAACTCTGGAA

TTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTGGATTATT

ATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTATCAGGGATC

TGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGATAAGTGTGAA

AACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGGGGGGCATATTA

ATGATGCCTTCA**TGA**CAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGTTGTTCTGCTTC

CTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATACCAAGAGCAGAT

CATATATTTTGTTTCACCATTCTTCTTTTGTAATAAATTTTGAATGTGCTTGAAAGTGAAAAG

CAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGACTCAAAATATTCTAA

AATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTGTAGTTATTGATTTAA

GCATTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTCAAAGACCTAAGGAAAA

ATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCATTGAAAATGGATCCTTTT

TGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAGTGAGAAGTAATTATTGTA

AATGGATGGATAAAAATGGAATTACTCATATACAGGGTGGAATTTTATCCTGTTATCACACCA

ACAGTTGATTATATATTTTCTGAATATCAGCCCCTAATAGGACAATTCTATTTGTTGACCATT

TCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAGTAATAATCATCTCTTTTTAAA

AAAAAAAAAAAAAAAAAAAAAA

FIGURE 482

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVAFSMRKVPN
REATEISHVLLCNVTQRVSFWFVVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLEFL
KIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKCENM
ITIENGIPSDPLDMKGGILMMPS

FIGURE 483

CGTCTCTGCGTTCGCC**ATG**CGTCCCGGGGCGCCAGGGCCACTCTGGCCTCTGCCCTGGGGGGC
CCTGGCTTGGGCCGTGGGCTTCGTGAGCTCCATGGGCTCGGGGAACCCCGCGCCCGGTGGTGT
TTGCTGGCTCCAGCAGGGCCAGGAGGCCACCTGCAGCCTGGTGCTCCAGACTGATGTCACCCG
GGCCGAGTGCTGTGCCTCCGGCAACATTGACACCGCCTGGTCCAACCTCACCCACCCGGGGAA
CAAGATCAACCTCCTCGGCTTCTTGGGCCTTGTCCACTGCCTTCCCTGCAAAGATTCGTGCGA
CGGCGTGGAGTGCGGCCCGGGCAAGGCGTGCCGCATGCTGGGGGGCCGCCCGCGCTGCGAGTG
CGCGCCCGACTGCTCGGGGCTCCCGGCGCGGCTGCAGGTCTGCGGCTCAGACGGCGCCACCTA
CCGCGACGAGTGCGAGCTGCGCGCCGCGCGCTGCCGCGGCCACCCGGACCTGAGCGTCATGTA
CCGGGGCCGCTGCCGCAAGTCCTGTGAGCACGTGGTGTGCCCGCGGCCACAGTCGTGCGTCGT
GGACCAGACGGGCAGCGCCCACTGCGTGGTGTGTCGAGCGGCGCCCTGCCCTGTGCCCTCCAG
CCCCGGCCAGGAGCTTTGCGGCAACAACAACGTCACCTACATCTCCTCGTGCCACATGCGCCA
GGCCACCTGCTTCCTGGGCCGCTCCATCGGCGTGCGCCACGCGGGCAGCTGCGCAGGCACCCC
TGAGGAGCCGCCAGGTGGTGAGTCTGCAGAAGAGGAAGAGAACTTCGTG**TGA**GCCTGCAGGAC
AGGCCTGGGCCTGGTGCCCGAGGCCCCCCATCATCCCCTGTTATTTATTGCCACAGCAGAGTC
TAATTTATATGCCACGGACACTCCTTAGAGCCCGGATTCGGACCACTTGGGGATCCCAGAACC
TCCCTGACGATATCCTGGAAGGACTGAGGAAGGGAGGCCTGGGGGCCGGCTGGTGGGTGGGAT
AGACCTGCGTTCCGGACACTGAGCGCCTGATTTAGGGCCCTTCTCTAGGATGCCCCAGCCCCT
ACCCTAAGACCTATTGCCGGGGAGGATTCCACACTTCCGCTCCTTTGGGGATAAACCTATTAA
TTATTGCTACTATCAAGAGGGCTGGGCATTCTCTGCTGGTAATTCCTGAAGAGGCATGACTGC
TTTTCTCAGCCCCAAGCCTCTAGTCTGGGTGTGTACGGAGGGTCTAGCCTGGGTGTGTACGGA
GGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGT
GAGTACGGAGGGTCTAGCCTGGGTGTGTATGGAGGATCTAGCCTGGGTGAGTATGGAGGGTCT
AGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTGGGTGAGTAT
GGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTG
GGTGTGTACGGAGGGTCTAGTCTGAGTGCGTGTGGGGACCTCAGAACACTGTGACCTTAGCCC
AGCAAGCCAGGCCCTTCATGAAGGCCAAGAAGGCTGCCACCATTCCCTGCCAGCCCAAGAACT
CCAGCTTCCCCACTGCCTCTGTGTGCCCCTTTGCGTCCTGTGAAGGCCATTGAGAAATGCCCA
GTGTGCCCCCTGGGAAAGGGCACGGCCTGTGCTCCTGACACGGGCTGTGCTTGGCCACAGAAC
CACCCAGCGTCTCCCCTGCTGCTGTCCACGTCAGTTCATGAGGCAACGTCGCGTGGTCTCAGA
CGTGGAGCAGCCAGCGGCAGCTCAGAGCAGGGCACTGTGTCCGGCGGAGCCAAGTCCACTCTG
GGGGAGCTCTGGCGGGGACCACGGGCCACTGCTCACCCACTGGCCCCGAGGGGGGTGTAGACG
CCAAGACTCACGCATGTGTGACATCCGGAGTCCTGGAGCCGGGTGTCCCAGTGGCACCACTAG
GTGCCTGCTGCCTCCACAGTGGGGTTCACACCCAGGGCTCCTTGGTCCCCCACAACCTGCCCC
GGCCAGGCCTGCAGACCCAGACTCCAGCCAGACCTGCCTCACCCACCAATGCAGCCGGGGCTG
GCGACACCAGCCAGGTGCTGGTCTTGGGCCAGTTCTCCCACGACGGCTCACCCTCCCCTCCAT
CTGCGTTGATGCTCAGAATCGCCTACCTGTGCCTGCGTGTAAACCACAGCCTCAGACCAGCTA
TGGGGAGAGGACAACACGGAGGATATCCAGCTTCCCCGGTCTGGGGTGAGGAATGTGGGGAGC
TTGGGCATCCTCCTCCAGCCTCCTCCAGCCCCCAGGCAGTGCCTTACCTGTGGTGCCCAGAAA
AGTGCCCCTAGGTTGGTGGGTCTACAGGAGCCTCAGCCAGGCAGCCCACCCCACCCTGGGGCC
CTGCCTCACCAAGGAAATAAAGACTCAAGCCATAAAAAAAA

FIGURE 484

MRPGAPGPLWPLPWGALAWAVGFVSSMGSGNPAPGGVCWLQQGQEATCSLVLQTDVTRAECCA

SGNIDTAWSNLTHPGNKINLLGFLGLVHCLPCKDSCDGVECGPGKACRMLGGRPRCECAPDCS

GLPARLQVCGSDGATYRDECELRAARCRGHPDLSVMYRGRCRKSCEHVVCPRPQSCVVDQTGS

AHCVVCRAAPCPVPSSPGQELCGNNNVTYISSCHMRQATCFLGRSIGVRHAGSCAGTPEEPPG

GESAEEEENFV

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 73-77, 215-219

Osteonectin domain proteins.

amino acids 97-130, 169-202

FIGURE 485

GCTCGAGGCCGGCGGCGGCGGGAGAGCGACCCGGGCGGCCTCGTAGCGGGGCCCCGGATCCCC

GAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAG**ATG**ATGGGCTTGGGA

AACGGGCGTCGCAGCATGAAGTCGCCGCCCCTCGTGCTGGCCGCCCTGGTGGCCTGCATCATC

GTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGATCATG

GAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGCGCCGTGGAGCTGAAGAAGAAC

GAGTTCCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGCCACAAC

TTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTGGTGAATAACATC

ACCACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGAGGAATTAC

GGCAGGCTGCAGCAGGATGTCCTCCAGTTTCAGAAGAACCAGACCAACCTGGAGAGGAAGTTC

TCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTGTGAGGAGCGA

ATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGAGTGAAAACAAC

GACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAGGCAGCAGGCCTG

CCACACACAGAGGTGCCACAAGGGAAGGGAAACGTGCTTGGTAACAGCAAGTCCCAGACACCA

GCCCCCAGTTCCGAAGTGGTTTTGGATTCAAAGAGACAAGTTGAGAAAGAGGAAACCAATGAG

ATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAGCCAGGCCGGGAGCAG

GTGGTGGAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGGAGAACTGGGCCAGACC

CCACAGGTGCAGGCTGCCCTGTCAGTGAGCCAGGAAAATCCAGAGATGGAGGGCCCTGAGCGA

GACCAGCTTGTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCTGCCGGGGAAGGGAGAAAC

CAGCAGAAACTGAGAGGAGAAGATGACTACAACATGGATGAAAATGAAGCAGAATCTGAGACA

GACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAACATAGATGTTTTTAATGTTGAAGATCAG

AAAAGAGACACCATAAATTTACTTGATCAGCGTGAAAAGCGGAATCATACACTC**TGA**ATTGAA

CTGGAATCACATATTTCACAACAGGGCCGAAGAGATGACTATAAAATGTTCATGAGGGACTGA

ATACTGAAAACTGTGAAATGTACTAAATAAAATGTACATCTGA

FIGURE 486

MMGLGNGRRSMKSPPLVLAALVACIIVLGFNYWIASSRSVDLQTRIMELEGRVRRAAAERGAV
ELKKNEFQGELEKQREQLDKIQSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQLKT
LQRNYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVASRD
LSENNDQRQQLQALSEPQPRLQAAGLPHTEVPQGKGNVLGNSKSQTPAPSSEVVLDSKRQVEK
EETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQENPEM
EGPERDQLVIPDGQEEEQEAAGEGRNQQKLRGEDDYNMDENEAESETDKQAALAGNDRNIDVF
NVEDQKRDTINLLDQREKRNHTL

Important features:
Signal peptide:
amino acids 1-29

FIGURE 487

AACTCAAACTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGGG

TGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATG**TAT

GGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATACAGCTCACAGCTCTTTGGCCT

ATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGCTCGG

TTAAAATGCACTTTCTCCAGCTTTGCCCCTGTGGGTGATGCTCTAACAGTGACCTGGAATTTT

CGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTCCAACCC

ATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGATGCCTCC

ATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGAAGAACCCA

CCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTACGCTTCTCT

GAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCATAATAGTAATT

GTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTCATAAAGTGGTG

GAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCTGTTTATTTAGAA

GACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAAGAACCCTAGTATT

TCTTGAAGTTAATGGAAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGTTTTCCAACCAGTTC

TGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCACAGTGCTCCTCCATAT

CACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGAGTGTAAATTTTTTCAA

GTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCTTAAGACACTACTTACAGTGT

TATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCCAATTTGTCTGTTACATTT

CCTTTCACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTAATGTGTTTACTCTCTTTCC

TTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTGTTTCTGATTAACAGTAAATC

CTAAATTCAAACTGTTAAATGACATTTTTATTTTTATGTCTCTCCTTAACTATGAGACACATC

TTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATTTTTGTCG

FIGURE 488

MYGKSSTRAVLLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVTW
NFRPLDGGPEQFVFYYHIDPFQPMSGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQVK
NPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLFQHYRKKRWAERAHK
VVEIKSKEEERLNQEKKVSVYLEDTD

FIGURE 489

AAGCAACCAAACTGCAAGCTTTGGGAGTTGTTCGCTGTCCCTGCCCTGCTCTGCTAGGGAGAG
AACGCCAGAGGGAGGCGGCTGGCCCGGCGGCAGGCTCTCAGAACCGCTACCGGCG**ATG**CTACT
GCTGTGGGTGTCGGTGGTCGCAGCCTTGGCGCTGGCGGTACTGGCCCCCGGAGCAGGGGAGCA
GAGGCGGAGAGCAGCCAAAGCGCCCAATGTGGTGCTGGTCGTGAGCGACTCCTTCGATGGAAG
GTTAACATTTCATCCAGGAAGTCAGGTAGTGAAACTTCCTTTTATCAACTTTATGAAGACACG
TGGGACTTCCTTTCTGAATGCCTACACAAACTCTCCAATTTGTTGCCCATCACGCGCAGCAAT
GTGGAGTGGCCTCTTCACTCACTTAACAGAATCTTGGAATAATTTTAAGGGTCTAGATCCAAA
TTATACAACATGGATGGATGTCATGGAGAGGCATGGCTACCGAACACAGAAATTTGGGAAACT
GGACTATACTTCAGGACATCACTCCATTAGTAATCGTGTGGAAGCGTGGACAAGAGATGTTGC
TTTCTTACTCAGACAAGAAGGCAGGCCCATGGTTAATCTTATCCGTAACAGGACTAAAGTCAG
AGTGATGGAAAGGGATTGGCAGAATACAGACAAAGCAGTAAACTGGTTAAGAAAGGAAGCAAT
TAATTACACTGAACCATTTGTTATTTACTTGGGATTAAATTTACCACACCCTTACCCTTCACC
ATCTTCTGGAGAAAATTTTGGATCTTCAACATTTCACACATCTCTTTATTGGCTTGAAAAAGT
GTCTCATGATGCCATCAAAATCCCAAAGTGGTCACCTTTGTCAGAAATGCACCCTGTAGATTA
TTACTCTTCTTATACAAAAAACTGCACTGGAAGATTTACAAAAAAAGAAATTAAGAATATTAG
AGCATTTTATTATGCTATGTGTGCTGAGACAGATGCCATGCTTGGTGAAATTATTTTGGCCCT
TCATCAATTAGATCTTCTTCAGAAAACTATTGTCATATACTCCTCAGACCATGGAGAGCTGGC
CATGGAACATCGACAGTTTTATAAAATGAGCATGTACGAGGCTAGTGCACATGTTCCGCTTTT
GATGATGGGACCAGGAATTAAAGCCGGCCTACAAGTATCAAATGTGGTTTCTCTTGTGGATAT
TTACCCTACCATGCTTGATATTGCTGGAATTCCTCTGCCTCAGAACCTGAGTGGATACTCTTT
GTTGCCGTTATCATCAGAAACATTTAAGAATGAACATAAAGTCAAAAACCTGCATCCACCCTG
GATTCTGAGTGAATTCCATGGATGTAATGTGAATGCCTCCACCTACATGCTTCGAACTAACCA
CTGGAAATATATAGCCTATTCGGATGGTGCATCAATATTGCCTCAACTCTTTGATCTTTCCTC
GGATCCAGATGAATTAACAAATGTTGCTGTAAAATTTCCAGAAATTACTTATTCTTTGGATCA
GAAGCTTCATTCCATTATAAACTACCCTAAAGTTTCTGCTTCTGTCCACCAGTATAATAAAGA
GCAGTTTATCAAGTGGAAACAAAGTATAGGACAGAATTATTCAAACGTTATAGCAAATCTTAG
GTGGCACCAAGACTGGCAGAAGGAACCAAGGAAGTATGAAAATGCAATTGATCAGTGGCTTAA
AACCCATATGAATCCAAGAGCAGTT**TGA**ACAAAAAGTTTAAAAATAGTGTTCTAGAGATACAT
ATAAATATATTACAAGATCATAATTATGTATTTTAAATGAAACAGTTTTAATAATTACCAAGT
TTTGGCCGGGCACAGTGGCTCACACCTGTAATCCCAGGACTTTGGGAGGCTGAGGAAAGCAGA
TCACAAGGTCAAGAGATTGAGACCATCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAA
TACAAAAATTAGCTGGGCGCGGTGGTGCACACCTATAGTCTCAGCTACTCAGAGGCTGAGGCA
GGAGGATCGCTTGAACCCGGGAGGCAGCAGTTGCAGTGAGCTGAGATTGCGCCACTGTACTCC
AGCCTGGCAACAGAGTGAGACTGTGTCGCAAAAAAATAAAATAAAATAATAATAATTACCAA
TTTTTCATTATTTTGTAAGAATGTAGTGTATTTTAAGATAAAATGCCAATGATTATAAAATCA
CATATTTTCAAAAATGGTTATTATTTAGGCCTTTGTACAATTTCTAACAATTTAGTGGAAGTA
TCAAAAGGATTGAAGCAAATACTGTAACAGTTATGTTCCTTTAAATAATAGAGAATATAAAAT
ATTGTAATAATATGTATCATAAAATAGTTGTATGTGAGCATTTGATGGTGAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAA

FIGURE 490

MLLLWVSVVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDSFDGRLTFHPGSQVVKLPFINFM
KTRGTSFLNAYTNSPICCPSRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQKF
GKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNWLRK
EAINYTEPFVIYLGLNLPHPYPSPSSGENFGSSTFHTSLYWLEKVSHDAIKIPKWSPLSEMHP
VDYYSSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQLDLLQKTIVIYSSDHG
ELAMEHRQFYKMSMYEASAHVPLLMMGPGIKAGLQVSNVVSLVDIYPTMLDIAGIPLPQNLSG
YSLLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYSDGASILPQLFD
LSSDPDELTNVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIGQNYSNVIA
NLRWHQDWQKEPRKYENAIDQWLKTHMNPRAV

Important features:
Signal peptide:
amino acids 1-15

N-glycosylation sites.
amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,
498-501

Sulfatases proteins:
amino acids 286-315, 359-369, 78-97

FIGURE 491

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGA

GCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**GC

CTCTCTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGT

TGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGT

TGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGA

CATCTATAGCACCCTTCTGGGCCTGCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGAC

ATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTT

CTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGG

AGGCCTCCTGGGATTCATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTC

ACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTC

TTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCG

CTCCAACTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGG

TCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGTATGTG**TGA**AGAAC

CAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCA

CAGGTGAGGGACACTACCACTGGATCGTGTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGG

CCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATGCAGGTTGAATTGCCAA

GGATGCTCGCCATGCCAGCCTTTCTGTTTTCCTCACCTTGCTGCTCCCCTGCCCTAAGTCCCC

AACCCTCAACTTGAAACCCATTCCCTTAAGCCAGGACTCAGAGGATCCCTTTGCCCTCTGGT

TTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCACTGACTGACCCTCTGTGATCAA

AGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGGGGATGGGAAGGAGAAGCAGT

GGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCCAAAGAAACTGATTGGCCC

TGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCAGACTAATTTGTGCATGAACTG

AAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATGCAGGATGGGAGGACAGGAA

GGCAGCCTGGGACATTTAAAAAAATA

FIGURE 492

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQ

CDIYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFI

LGGLLGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQR

NRSNYYDAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 493

GCACTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCC**ATG**GTGCCCAGAATCTTCGCTCCTGCT

TATGTGTCAGTCTGTCTCCTCCTCTTGTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAGAA

CCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAGTGC

TGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCCTGCACCTTC

TGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCCTTTGGCCTCACAAACGATTTTGTTGTGAAG

CTGAAGGTTCAGGGTGTGAATTCCCAGTGCCACTCATCTCCCATCTCCAGTAAATGTGAAAGC

AGAAGACGTTTTCCC**TGA**GAAGACATAGAAAGAAAATCAACTTTCACTAAGGCATCTCAGAAA

CATAGGCTAAGGTAATATGTGTACCAGTAGAGAAGCCTGAGGAATTTACAAAATGATGCAGCT

CCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGTAGATTATCA

GGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

FIGURE 494

MVPRIFAPAYVSVCLLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYNDAIVSLSET
RQCGPPCTFWPCFELCCLDSFGLTNDFVVKLKVQGVNSQCHSSPISSKCESRRRFP

Important features:

Signal peptide:

amino acids 1-25

FIGURE 495

CTCCACTGCAACCACCCAGAGCC**ATG**GCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATTT

TCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCTGA

TGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCTGCAGCACTGTTGCT

ATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAAACTGCACCTTCAGAGTCT

GCTTTGAGCAGTGCTGCCCCTGGACCTTCATGGTGAAGCTGATAAACCAGAACTGCGACTCAG

CCCGGACCTCGGATGACAGGCTTTGTCGCAGTGTCAGC**TAA**TGGAACATCAGGGGAACGATGA

CTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTTACCTGAGATCTGGGA

TGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCACTTCATTCTGTGACC

TGTCTGAGGCCCACCCTGCAGCTGCCCTGAGGAGGCCCACAGGTCCCCTTCTAGAATTCTGGA

CAGCATGAGATGCGTGTGCTGATGGGGGCCCAGGGACTCTGAACCCTCCTGATGACCCCTATG

GCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCTTCACCCTTCTGTGAGATT

TTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAATAAATTTATGTA

CTTTATAAATGAAAA

FIGURE 496

MAPRGCIVAVFAIFCISRLLCSHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVPL

ARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCDSARTSDDRLCRSVS

Important features:

Signal peptide:

amino acids 1-24

FIGURE 497

```
TGAAGGACTTTTCCAGGACCCAAGGCCACACACTGGAAGTCTTGCAGCTGAAGGGAGGCACTC
CTTGGCCTCCGCAGCCGATCACATGAAGGTGGTGCCAAGTCTCCTGCTCTCCGTCCTCCTGGC
ACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCCAGCCCCTCA
GAACCAGACCAGCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGGCCAG
CGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGGCAGCAGCTTGCCAA
GGAGACTTCAAACTTCGGATTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGGCAACAT
GGTCTTCTCTCCATTTGGCATGTCCTTGGCCATGACAGGCTTGATGCTGGGGGCCACAGGGCC
GACTGAAACCCAGATCAAGAGAGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAGCCCGGGCT
CCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACTGGGCCTCTC
ACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCTTCAATTTATC
CAAGAGGTATTTTGATACAGAGTGCGTGCCTATGAATTTTCGCAATGCCTCACAGGCCAAAAG
GCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTCCCAAACTGTTTGATGAGAT
TAATCCTGAAACCAAATTAATTCTTGTGGATTACATCTTGTTCAAAGGGAAATGGTTGACCCC
ATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTACAAGACCATTAAGGT
GCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAATTTTCGTTGTCATGT
CCTCAAACTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCCTCATGGAGAAAATGGGTGA
CCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACATGGCTCAGAAACATGAA
AACCAGAAACATGGAAGTTTTCTTTCCGAAGTTCAAGCTAGATCAGAAGTATGAGATGCATGA
GCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCTTTGCTGACCTTAGTGAACTCTC
AGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGAACAGTGATTGAAGTTGATGA
AAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTACTGCTTATTCCATGCCTCCTGT
CATCAAAGTGGACCGGCCATTTCATTTCATGATCTATGAAGAAACCTCTGGAATGCTTCTGTT
TCTGGGCAGGGTGGTGAATCCGACTCTCCTATAATTCAGGACATGCATAAGCACTTCGTGCTG
TAGTAGATGCTGAATCTGAGGTATCAAACACACACAGGATACCAGCAATGGATGGCAGGGGAG
AGTGTTCCTTTTGTTCTTAACTAGTTTAGGGTGTTCTCAAATAAATACAGTAGTCCCCACTTA
TCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGCCTGAAACGGTGGACAGTGCTGAACCT
TATATATATTTTTTCCTACACATACATACCTATGATAAAGTTTAATTTATAAATTAGGCACAG
TAAGAGATTAACAATAATAACAACATTAAGTAAAATGAGTTACTTGAACGCAAGCACTGCAAT
ACCATAACAGTCAAACTGATTATAGAGAAGGCTACTAAGTGACTCATGGGCGAGGAGCATAGA
CAGTGTGGAGACATTGGGCAAGGGGAGAATTCACATCCTGGGTGGGACAGAGCAGGACGATGC
AAGATTCCATCCCACTACTCAGAATGGCATGCTGCTTAAGACTTTTAGATTGTTTATTTCTGG
AATTTTTCATTTAATGTTTTTGGACCATGGTTGACCATGGTTAACTGAGACTGCAGAAGCAA
AACCATGGATAAGGGAGGACTACTACAAAGCATTAAATTGATACATATTTTTTAAAAAAAAA
AAAAAAAAA
```

FIGURE 498

MKVVPSLLLSVLLAQVWLVPGLAPSPQSPETPAPQNQTSRVVQAPREEEEDEQEASEEKAGEE
EKAWLMASRQQLAKETSNFGFSLLRKISMRHDGNMVFSPFGMSLAMTGLMLGATGPTETQIKR
GLHLQALKPTKPGLLPSLFKGLRETLSRNLELGLSQGSFAFIHKDFDVKETFFNLSKRYFDTE
CVPMNFRNASQAKRLMNHYINKETRGKIPKLFDEINPETKLILVDYILFKGKWLTPFDPVFTE
VDTFHLDKYKTIKVPMMYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHLALEDY
LTTDLVETWLRNMKTRNMEVFFPKFKLDQKYEMHELLRQMGIRRIFSPFADLSELSATGRNLQ
VSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLFLGRVVNP
TLL

FIGURE 499

CTAGCCTGCGCCAAGGGGTAGTGAGACCGCGCGGCAACAGCTTGCGGCTGCGGGGAGCTCCCG

TGGGCGCTCCGCTGGCTGTGCAGGCGGCC**ATG**GATTCCTTGCGGAAAATGCTGATCTCAGTCG

CAATGCTGGGCGCAGGGGCTGGCGTGGGCTACGCGCTCCTCGTTATCGTGACCCCGGGAGAGC

GGCGGAAGCAGGAAATGCTAAAGGAGATGCCACTGCAGGACCCAAGGAGCAGGGAGGAGGCGG

CCAGGACCCAGCAGCTATTGCTGGCCACTCTGCAGGAGGCAGCGACCACGCAGGAGAACGTGG

CCTGGAGGAAGAACTGGATGGTTGGCGGCGAAGGCGGCGCCAGCGGGAGGTCACCG**TGA**GACC

GGACTTGCCTCCGTGGGCGCCGGACCTTGGCTTGGGCGCAGGAATCCGAGGCAGCCTTTCTCC

TTCGTGGGCCCAGCGGAGAGTCCGGACCGAGATACCATGCCAGGACTCTCCGGGGTCCTGTGA

GCTGCCGTCGGGTGAGCACGTTTCCCCCAAACCCTGGACTGACTGCTTTAAGGTCCGCAAGGC

GGGCCAGGGCCGAGACGCGAGTCGGATGTGGTGAACTGAAAGAACCAATAAAATCATGTTCCT

CCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 500

MDSLRKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLAT

LQEAATTQENVAWRKNWMVGGEGGASGRSP

Important features:

Signal peptide:

amino acids 1-18

FIGURE 501

CAGGAGAGAAGGCACCGCCCCCACCCCGCCTCCAAAGCTAACCCTCGGGCTTGAGGGGAAGAG
GCTGACTGTACGTTCCTTCTACTCTGGCACCACTCTCCAGGCTGCC**ATG**GGGCCCAGCACCCC
TCTCCTCATCTTGTTCCTTTTGTCATGGTCGGGACCCCTCCAAGGACAGCAGCACCACCTTGT
GGAGTACATGGAACGCCGACTAGCTGCTTTAGAGGAACGGCTGGCCCAGTGCCAGGACCAGAG
TAGTCGGCATGCTGCTGAGCTGCGGGACTTCAAGAACAAGATGCTGCCACTGCTGGAGGTGGC
AGAGAAGGAGCGGGAGGCACTCAGAACTGAGGCCGACACCATCTCCGGGAGAGTGGATCGTCT
GGAGCGGGAGGTAGACTATCTGGAGACCCAGAACCCAGCTCTGCCCTGTGTAGAGTTTGATGA
GAAGGTGACTGGAGGCCCTGGGACCAAAGGCAAGGGAAGAAGGAATGAGAAGTACGATATGGT
GACAGACTGTGGCTACACAATCTCTCAAGTGAGATCAATGAAGATTCTGAAGCGATTTGGTGG
CCCAGCTGGTCTATGGACCAAGGATCCACTGGGGCAAACAGAGAAGATCTACGTGTTAGATGG
GACACAGAATGACACAGCCTTTGTCTTCCCAAGGCTGCGTGACTTCACCCTTGCCATGGCTGC
CCGGAAAGCTTCCCGAGTCCGGGTGCCCTTCCCCTGGGTAGGCACAGGGCAGCTGGTATATGG
TGGCTTTCTTTATTTTGCTCGGAGGCCTCCTGGAAGACCTGGTGGAGGTGGTGAGATGGAGAA
CACTTTGCAGCTAATCAAATTCCACCTGGCAAACCGAACAGTGGTGGACAGCTCAGTATTCCC
AGCAGAGGGGCTGATCCCCCCCTACGGCTTGACAGCAGACACCTACATCGACCTGGTAGCTGA
TGAGGAAGGTCTTTGGGCTGTCTATGCCACCCGGGAGGATGACAGGCACTTGTGTCTGGCCAA
GTTAGATCCACAGACACTGGACACAGAGCAGCAGTGGGACACACCATGTCCCAGAGAGAATGC
TGAGGCTGCCTTTGTCATCTGTGGGACCCTCTATGTCGTCTATAACACCCGTCCTGCCAGTCG
GGCCCGCATCCAGTGCTCCTTTGATGCCAGCGGCACCCTGACCCCTGAACGGGCAGCACTCCC
TTATTTTCCCCGCAGATATGGTGCCCATGCCAGCCTCCGCTATAACCCCCGAGAACGCCAGCT
CTATGCCTGGGATGATGGCTACCAGATTGTCTATAAGCTGGAGATGAGGAAGAAAGAGGAGGA
GGTT**TGA**GGAGCTAGCCTTGTTTTTTGCATCTTTCTCACTCCCATACATTTATATTATATCCC
CACTAAATTTCTTGTTCCTCATTCTTCAAATGTGGGCCAGTTGTGGCTCAAATCCTCTATATT
TTTAGCCAATGGCAATCAAATTCTTTCAGCTCCTTTGTTTCATACGGAACTCCAGATCCTGAG
TAATCCTTTTAGAGCCCGAAGAGTCAAAACCCTCAATGTTCCCTCCTGCTCTCCTGCCCCATG
TCAACAAATTTCAGGCTAAGGATGCCCCAGACCCAGGGCTCTAACCTTGTATGCGGGCAGGCC
CAGGGAGCAGGCAGCAGTGTTCTTCCCCTCAGAGTGACTTGGGGAGGGAGAAATAGGAGGAGA
CGTCCAGCTCTGTCCTCTCTTCCTCACTCCTCCCTTCAGTGTCCTGAGGAACAGGACTTTCTC
CACATTGTTTTGTATTGCAACATTTTGCATTAAAAGGAAAATCCACAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 502

MGPSTPLLILFLLSWSGPLQGQQHHLVEYMERRLAALEERLAQCQDQSSRHAAELRDFKNKML
PLLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGRRN
EKYDMVTDCGYTISQVRSMKILKRFGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRLRDF
TLAMAARKASRVRVPFPWVGTGQLVYGGFLYFARRPPGRPGGGGEMENTLQLIKFHLANRTVV
DSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCLAKLDPQTLDTEQQWDTP
CPRENAEAAFVICGTLYVVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAHASLRYN
PRERQLYAWDDGYQIVYKLEMRKKEEEV

Important features:

Signal peptide:
amino acids 1-21

N-glycosylation sites.
amino acids 177-180, 248-251

FIGURE 503

TGCGGCGCAGTGTAGACCTGGGAGG**ATG**GGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGGT

CTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGGCC

CTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAAGAA

CGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCCTCTCAGCA

CGGGCTGGGAGGGTGTGACCAGAGTGTCATGGACCTGATAAAGCGAAACTCCGGATGGGTGTT

TGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGACTATGC

CATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGTACAGTCT

GACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTCACCAAGTGGAGCAGGAGCCTGGGCTT

CCTGTCACAG**TAG**CAGGCCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAGATCCTTCTGTG

AGTGCTGCGTCCCCAGTAGGGATGGCGCCCACAGGGTCCTGTGACCTCGGCCAGTGTCCACCC

ACCTCGCTCAGCGGCTCCCGGGGCCCAGCACCAGCTCAGAATAAAGCGATTCCACAGCA

FIGURE 504

MGGLLLAAFLALVSVPRAQAVWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVTL
TPENNLRTLSSQHGLGGCDQSVMDLIKRNSGWVFENPSIGVLELWVLATNFRDYAIIFTQLEF
GDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

Important features:

Signal peptide:

amino acids 1-20

FIGURE 505

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCACA

GCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCCTCCAGGCC**ATG**AGGATTCTGCAGTTAATC

CTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAGTGC

AAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTCGAGAAGACGCGGCTACTCTGTGGGGCG

ACGCTCATCGCCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACATAGTT

CACCTGGGGCAGCACAACCTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCCACTGAG

TCCTTCCCCCACCCCGGCTTCAACAACAGCCTCCCCAACAAAGACCACCGCAATGACATCATG

CTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCCTCTCCTCA

CGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCCAGCCCCCAG

TTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTGAGCACCAGAAGTGTGAG

AACGCCTACCCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGGAAGGGGGCAAG

GACTCCTGCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCAGTCTCTTCAAGGCATTATC

TCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACACGAAAGTCTGCAAA

TATGTGGACTGGATCCAGGAGACGATGAAGAACAAT**TAG**ACTGGACCCACCCACCACAGCCCA

TCACCCTCCATTTCCACTTGGTGTTTGGTTCCTGTTCACTCTGTTAATAAGAAACCCTAAGCC

AAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATGCTGTCACTTAATAATC

AACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTTGAAATATTGTGACTCTG

GGAATGACAACACCTGGTTTGTTCTCTGTTGTATCCCCAGCCCCAAAGACAGCTCCTGGCCAT

ATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAA

FIGURE 506

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAHC
LKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITWAV
RPLTLSSRCVTAGTSCLISGWGSTSSPQLRLPHTLRCANITIIEHQKCENAYPGNITDTMVCA
SVQEGGKDSCQGDSGGPLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIQETMKNN

Important features:

Signal peptide:

amino acids 1-18

Serine proteases, trypsin family, histidine active site.

amino acids 58-63

N-glycosylation sites.

amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.

amino acids 145-148

Kringle domain proteins.

amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein

amino acids 199-209, 47-63, 220-243

Apple domain proteins

amino acids 222-249, 189-222

FIGURE 507

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGAC**ATG**AGG

AGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAGCA

CCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAGGCC

TGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGCTGTTCCCTGTC

CAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTCAGGGCAGGGGCCCCATCCTT

CCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCCTGAGTCCCGAGCCC

GACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCCCCGGTTG

TGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACATCTACCAC

CCCCAG**TAG**GGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGCCCAGGCTGTTGGGA

CTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCCAGCAGGCAAAAAAAAAAAAA

AAAAAA

FIGURE 508

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVLF
PVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEERP
RLWVMPNHQVLLGPEEDQDHIYHPQ

FIGURE 509

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCCC

ACGAGGCTGCCGCATCCTGCCCTCGGAACA**ATG**GGACTCGGCGCGCGAGGTGCTTGGGCCGCG

CTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCCATG

GCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACTCCAGTGCTAACTCAACAGAG

ACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAACCACCA

ACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCGGCATCT

AATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTCTACACCC

AAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCGTAACCCAC

AATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACTATGCATTCTGAAGCA

AAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAACGCTGGGAGTT

TTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTCGGTATCGAACC

ATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGAATACAGATTGAT

GCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAACAATATTCTCTTTTTGAAAATA

GTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTAAAGATTCTTCAAGG

TAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGTTCATACAATGGTTTT

AGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTGGCTGGGGTGGGGGCATTGG

TCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAATGCCATCTGGGCATACA

AATAAGAAGTTTGTCACAGCACTCAGGATTTTGGGTATCTTTTGTAGCTCACATAAAGAACTT

CAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACACAGAATTATACAATCAAA

CTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTGTGCTTTAAACTGTAGTAGTT

GGTCTAGAAACAAAATACTCC

FIGURE 510

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPSDHT
NETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQNTS
QISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSILYIGCKM
YYSRRGIRYRTIDEHDAII

FIGURE 511

GACTTTGCTTGAATGTTTACATTTTCTGCTCGCTGTCCTACATATCACAATATAGTGTTCACGTTTTGTTAAAAC

TTTGGGGTGTCAGGAGTTGAGCTTGCTCAGCAAGCCAGC**ATG**GCTAGGATGAGCTTTGTTATAGCAGCTTGCCAA

TTGGTGCTGGGCCTACTAATGACTTCATTAACCGAGTCTTCCATACAGAATAGTGAGTGTCCACAACTTTGCGTA

TGTGAAATTCGTCCCTGGTTTACCCCACAGTCAACTTACAGAGAAGCCACCACTGTTGATTGCAATGACCTCCGC

TTAACAAGGATTCCCAGTAACCTCTCTAGTGACACACAAGTGCTTCTCTTACAGAGCAATAACATCGCGAAGACT

GTGGATGAGCTGCAGCAGCTTTTCAACTTGACTGAACTAGATTTCTCCCAAAACAACTTTACTAACATTAAGGAG

GTCGGGCTGGCAAACCTAACCCAGCTCACAACGCTGCATTTGGAGGAAAATCAGATTACCGAGATGACTGATTAC

TGTCTACAAGACCTCAGCAACCTTCAAGAACTCTACATCAACCACAACCAAATTAGCACTATTTCTGCTCATGCT

TTTGCAGGCTTAAAAAATCTATTAAGGCTCCACCTGAACTCCAACAAATTGAAAGTTATTGATAGTCGCTGGTTT

GATTCTACACCCAACCTGGAAATTCTCATGATCGGAGAAAACCCTGTGATTGGAATTCTGGATATGAACTTCAAA

CCCCTCGCAAATTTGAGAAGCTTAGTTTTGGCAGGAATGTATCTCACTGATATTCCTGGAAATGCTTTGGTGGGT

CTGGATAGCCTTGAGAGCCTGTCTTTTTATGATAACAAACTGGTTAAAGTCCCTCAACTTGCCCTGCAAAAAGTT

CCAAATTTGAAATTCTTAGACCTCAACAAAAACCCCATTCACAAAATCCAAGAAGGGGACTTCAAAAATATGCTT

CGGTTAAAAGAACTGGGAATCAACAATATGGGCGAGCTCGTTTCTGTCGACCGCTATGCCCTGGATAACTTGCCT

GAACTCACAAAGCTGGAAGCCACCAATAACCCTAAACTCTCTTACATCCACCGCTTGGCTTTCCGAAGTGTCCCT

GCTCTGGAAAGCTTGATGCTGAACAACAATGCCTTGAATGCCATTTACCAAAAGACAGTCGAATCCCTCCCCAAT

CTGCGTGAGATCAGTATCCATAGCAATCCCCTCAGGTGTGACTGTGTGATCCACTGGATTAACTCCAACAAAACC

AACATCCGCTTCATGGAGCCCCTGTCCATGTTCTGTGCCATGCCGCCCGAATATAAAGGGCACCAGGTGAAGGAA

GTTTTAATCCAGGATTCGAGTGAACAGTGCCTCCCAATGATATCTCACGACAGCTTCCCAAATCGTTTAAACGTG

GATATCGGCACGACGGTTTTCCTAGACTGTCGAGCCATGGCTGAGCCAGAACCTGAAATTTACTGGGTCACTCCC

ATTGGAAATAAGATAACTGTGGAAACCCTTTCAGATAAATACAAGCTAAGTAGCGAAGGTACCTTGGAAATATCT

AACATACAAATTGAAGACTCAGGAAGATACACATGTGTTGCCCAGAATGTCCAAGGGGCAGACACTCGGGTGGCA

ACAATTAAGGTTAACGGGACCCTTCTGGATGGTACCCAGGTGCTAAAAATATACGTCAAGCAGACAGAATCCCAT

TCCATCTTAGTGTCCTGGAAAGTTAATTCCAATGTCATGACGTCAAACTTAAAATGGTCGTCTGCCACCATGAAG

ATTGATAACCCTCACATAACATATACTGCCAGGGTCCCAGTCGATGTCCATGAATACAACCTAACGCATCTGCAG

CCTTCCACAGATTATGAAGTGTGTCTCACAGTGTCCAATATTCATCAGCAGACTCAAAAGTCATGCGTAAATGTC

ACAACCAAAAATGCCGCCTTCGCAGTGGACATCTCTGATCAAGAAACCAGTACAGCCCTTGCTGCAGTAATGGGG

TCTATGTTTGCCGTCATTAGCCTTGCGTCCATTGCTGTGTACTTTGCCAAAAGATTTAAGAGAAAAAACTACCAC

CACTCATTAAAAAAGTATATGCAAAAAACCTCTTCAATCCCACTAAATGAGCTGTACCCACCACTCATTAACCTC

TGGGAAGGTGACAGCGAGAAAGACAAAGATGGTTCTGCAGACACCAAGCCAACCCAGGTCGACACATCCAGAAGC

TATTACATGTGG**TAA**CTCAGAGGATATTTTGCTTCTGGTAGTAAGGAGCACAAAGACGTTTTTGCTTTATTCTGC

AAAAGTGAACAAGTTGAAGACTTTTGTATTTTTGACTTTGCTAGTTTGTGGCAGAGTGGAGAGGACGGGTGGATA

TTTCAAATTTTTTTAGTATAGCGTATCGCAAGGGTTTGACACGGCTGCCAGCGACTCTAGGCTTCCAGTCTGTGT

TTGGTTTTTATTCTTATCATTATTATGATTGTTATTATATTATTATTTTATTTTAGTTGTTGTGCTAAACTCAAT

AATGCTGTTCTAACTACAGTGCTCAATAAAATGATTAATGACAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 512

MARMSFVIAACQLVLGLLMTSLTESSIQNSECPQLCVCEIRPWFTPQSTYREATTVDCNDLRL
TRIPSNLSSDTQVLLLQSNNIAKTVDELQQLFNLTELDFSQNNFTNIKEVGLANLTQLTTLHL
EENQITEMTDYCLQDLSNLQELYINHNQISTISAHAFAGLKNLLRLHLNSNKLKVIDSRWFDS
TPNLEILMIGENPVIGILDMNFKPLANLRSLVLAGMYLTDIPGNALVGLDSLESLSFYDNKLV
KVPQLALQKVPNLKFLDLNKNPIHKIQEGDFKNMLRLKELGINNMGELVSVDRYALDNLPELT
KLEATNNPKLSYIHRLAFRSVPALESLMLNNNALNAIYQKTVESLPNLREISIHSNPLRCDCV
IHWINSNKTNIRFMEPLSMFCAMPPEYKGHQVKEVLIQDSSEQCLPMISHDSFPNRLNVDIGT
TVFLDCRAMAEPEPEIYWVTPIGNKITVETLSDKYKLSSEGTLEISNIQIEDSGRYTCVAQNV
QGADTRVATIKVNGTLLDGTQVLKIYVKQTESHSILVSWKVNSNVMTSNLKWSSATMKIDNPH
ITYTARVPVDVHEYNLTHLQPSTDYEVCLTVSNIHQQTQKSCVNVTTKNAAFAVDISDQETST
ALAAVMGSMFAVISLASIAVYFAKRFKRKNYHHSLKKYMQKTSSIPLNELYPPLINLWEGDSE
KDKDGSADTKPTQVDTSRSYYMW

Important features:
Signal peptide:
Amino acids 1-25

Transmembrane domain:
Amino acids 508-530

N-glycosylation sites:
Amino acids 69-73;96-100;106-110;117-121;385-389;517-521;
582-586;611-615

Tyrosine kinase phosphorylation site:
Amino acids 573-582

N-myristoylation sites:
Amino acids 16-22;224-230;464-470;637-643;698-704

FIGURE 513

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTCC

CAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCCCT

GTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAACCGAGGGCT

GAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGAATCACGCATGC

CGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGCAAGGA

GTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGATCAACCA

TGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACGCTGCTGG

ACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCACCAGGCTGG

GAAGGAAGCAGAGAAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGGAAAGGAAGT

GGAGAAGCTTGGCCAAGGTGCCCACCATGCTGCTGGCCAGGCCGGGAAGGAGCTGCAGAATGC

TCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGCAACCATCAAAG

CGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGGGGCCTCAGTCAA

CACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCATGCCC**TAA**ACTGG

CATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGACATGACCTGGAGGGG

TTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTTGTACTGGGATTTGTGAATAAA

CTTGATACACCA

FIGURE 514

MHLARLVGSCSLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAGR
EVEKVFNGLSNMGSHTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAGQA
GKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQNAHN
GVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP

Important features:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

FIGURE 515

CCCACGCGTCCGCCCACGCGTCCGGGTGCCACTCGCGCGCCGGCCGCGCTCCGGGCTTCTCTT

TTCCCTCCGACGCGCCACGGCTGCCCAGACATTCCGGCTGCCGGGTCTGGAGAGCTCCCCGAA

CCCCTCCGCGGAGAGGAGCGAGGCGGCGCCAGGGTGGCCCCCGGGGCGCGCTTGGTCTCGGAG

AAGCGGGGACGAGGCCGGAGGATGAGCGACTGAGGGCGACGCGGGCACTGACGCGAGTTGGGG

CCGCGACTACCGGCAGCTGACAGCGCGATGAGCGACTCCCCAGAGACGCCCTAGCCCGGTGTG

CGCGCCAGGCGGAGCGCGCAGGTGGGGCTGGGCTGTTAGTGGTCCGCCCCACGCGGGTCGCCG

GCCGGCCCAGGATGGGCGCTGGCAACCCGGGCCCGCGCCCGCCGCTGCTACCCCTGCGCCCGC

TGCGAGCCCGGCGTCCGGCCCGCGCCCTGCGCTCATGGACGGCGGCTCCCGGCTGGCGGCGGC

GCGCCCCGGGCTGTGAATGCGACTCGCCCCTCGGCCGCGCTCCCCGCCCGCCCGCCCGCCGG

GACGTGGTAGGGG**ATG**CCCAGCTCCACTGCGATGGCAGTTGGCGCGCTCTCCAGTTCCCTCCT

GGTCACCTGCTGCCTGATGGTGGCTCTGTGCAGTCCGAGCATCCCGCTGGAGAAGCTGGCCCA

GGCACCAGAGCAGCCGGGCCAGGAGAAGCGTGAGCACGCCACTCGGGACGGCCCGGGGCGGGT

GAACGAGCTCGGGCGCCCGGCGAGGGACGAGGGCGGCAGCGGCCGGGACTGGAAGAGCAAGAG

CGGCCGTGGGCTCGCCGGCCGTGAGCCGTGGAGCAAGCTGAAGCAGGCCTGGGTCTCCCAGGG

CGGGGGCGCCAAGGCCGGGGATCTGCAGGTCCGGCCCCGCGGGGACACCCCGCAGGCGGAAGC

CCTGGCCGCAGCCGCCCAGGACGCGATTGGCCCGGAACTCGCGCCCACGCCCGAGCCACCCGA

GGAGTACGTGTACCCGGACTACCGTGGCAAGGGCTGCGTGGACGAGAGCGGCTTCGTGTACGC

GATCGGGGAGAAGTTCGCGCCGGGCCCCTCGGCCTGCCCGTGCCTGTGCACCGAGGAGGGGCC

GCTGTGCGCGCAGCCCGAGTGCCCGAGGCTGCACCCGCGCTGCATCCACGTCGACACGAGCCA

GTGCTGCCCGCAGTGCAAGGAGAGGAAGAACTACTGCGAGTTCCGGGGCAAGACCTATCAGAC

TTTGGAGGAGTTCGTGGTGTCTCCATGCGAGAGGTGTCGCTGTGAAGCCAACGGTGAGGTGCT

ATGCACAGTGTCAGCGTGTCCCCAGACGGAGTGTGTGGACCCTGTGTACGAGCCTGATCAGTG

CTGTCCCATCTGCAAAAATGGTCCAAACTGCTTTGCAGAAACCGCGGTGATCCCTGCTGGCAG

AGAAGTGAAGACTGACGAGTGCACCATATGCCACTGTACTTATGAGGAAGGCACATGGAGAAT

CGAGCGGCAGGCCATGTGCACGAGACATGAATGCAGGCAAATG**TAG**ACGCTTCCCAGAACACA

AACTCTGACTTTTTCTAGAACATTTTACTGATGTGAACATTCTAGATGACTCTGGGAACTATC

AGTCAAAGAAGACTTTTGATGAGGAATAATGGAAAATTGTTGGTACTTTTCCTTTTCTTGATA

ACAGTTACTACAACAGAAGGAAATGGATATATTTCAAAACATCAACAAGAACTTTGGGCATAA

AATCCTTCTCTAAATAAATGTGCTATTTTCACAGTAAGTACACAAAGTACACTATTATATAT

CAAATGTATTTCTATAATCCCTCCATTAGAGAGCTTATATAAGTGTTTTCTATAGATGCAGAT

TAAAAATGCTGTGTTGTCAACCGTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 516

MPSSTAMAVGALSSSLLVTCCLMVALCSPSIPLEKLAQAPEQPGQEKREHATRDGPGRVNELG
RPARDEGGSGRDWKSKSGRGLAGREPWSKLKQAWVSQGGGAKAGDLQVRPRGDTPQAEALAAA
AQDAIGPELAPTPEPPEEYVYPDYRGKGCVDESGFVYAIGEKFAPGPSACPCLCTEEGPLCAQ
PECPRLHPRCIHVDTSQCCPQCKERKNYCEFRGKTYQTLEEFVVSPCERCRCEANGEVLCTVS
ACPQTECVDPVYEPDQCCPICKNGPNCFAETAVIPAGREVKTDECTICHCTYEEGTWRIERQA
MCTRHECRQM

Important features:

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 11-30

Glycosaminoglycan attachment site.

amino acids 80-83

N-myristoylation sites.

amino acids 10-15, 102-107, 103-108

Cell attachment sequence.

amino acids 114-117

EGF-like domain cysteine pattern signature.

amino acids 176-187

FIGURE 517

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCCT

TTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCTAC

GGACGACGCCT**ATG**AAGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTGTGC

CGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAAACTTGAATCATTATATACAAG

TTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAATCTAACT

CTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAAATTTAAGGAGCTAGTTACACATG

GAGACGCTTCAACTGAGAATGATGTTTTAACCAATCCTATCAGTGAAGAAACTACAACTTTCC

CTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAAACACACGGAAAGTACCCCATTCTGGT

CGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTGAAAATGAAG

AGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAAACTGAGGCACCAAGAATGTTGCCAGTTG

TTACTGAATCATCTACAAGTCCATATGTTACCTCATACAAGTCACCTGTCACCACTTTAGATA

AGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCTCAGCTCTCAGGTGAAACTG

CGATAGAAAAACCCGAAGAGTTTGGAAAGCACCCAGAGAGTTGGAATAATGATGACATTTTGA

AAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACTTCTTAGTGACACCAGCAACCCAG

CATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCCTTGCTCTAGCAGCAG

CAGCAGAACATAAATTAAAAACAATGTATAAGTCCCAGTTATTGCCAGTAGGACGAACAAGTA

ATAAAATTGATGACATCGAAACTGTTATTAACATGCTGTGTAATTCTAGATCTAAACTCTATG

AATATTTAGATATTAAATGTGTTCCACCAGAGATGAGAGAAAAAGCTGCTACAGTATTCAATA

CATTAAAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTAAAAGTTTAT**TAA**ACAATAA

TATAAAAATTTTAAACCTACTTGATATTCCATAACAAAGCTGATTTAAGCAAACTGCATTTTT

TCACAGGAGAAATAATCATATTCGTAATTTCAAAAGTTGTATAAAAATATTTTCTATTGTAGT

TCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACAATTTTCATATGCACTAAAAACC

TAATTTAAAATAAAATTTTGGTTCAGGAAAAAA

FIGURE 518

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRSVPSGEPGREKKSNSPKH
VYSIASKGSKFKELVTHGDASTENDVLTNPISEETTTFPTGGFTPEIGKKKHTESTPFWSIKP
NNVSIVLHAEEPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSTSPYVTSYKSPVTTLDKSTG
IEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNPAYRE
DIEASKDHLKRSLALAAAAEHKLKTMYKSQLLPVGRTSNKIDDIETVINMLCNSRSKLYEYLD
IKCVPPEMREKAATVFNTLKNMCRSRRVTALLKVY

Important features:

Signal peptide:

amino acids 1-19

FIGURE 519

CGGCTCGAGTGCAGCTGTGGGGAGATTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTTG

GATTTGAAAGTTGAGAGCAGC**ATG**TTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTACTG

GATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTGGGT

GATTCAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGATAGAC

TGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCAATCTC

AGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGCAATGAT

GGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGAAATCCGC

CTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAGAGGAGCCC

AAAGAGCTCATGGTCCATGTGGGTGGATTGATTCAGATGGGATGTGTTTTCCAGAGCACAGAA

GTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGAGGAGATTGTA

TTTCGTTACTACCACAAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGGGCCACTTCCAG

AATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATGCTTCAAGGAGTG

AGGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCTGGTGTTCAAGAAA

ACCATTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCCCGGCAGCCCTGAGG

CCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTCTGTGCCACAATCCTG

CTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAAGAGTTCAGTGAATTCT

ACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAGAAAAACCCTGCCATTTT

GAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTACGGGAGGTGATCGAGGAA

GAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCACCCAGTTTGGCCTTCTCTG

AGGTCAGATCGGAACAACTCACTTGAAAAAAAGTCAGGTGGGGGAATGCCAAAAACACAGCAA

GCCTTT**TGA**GAAGAATGGAGAGTCCCTTCATCTCAGCAGCGGTGGAGACTCTCTCCTGTGTGT

GTCCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTCCCAGCTGTCCTCCTGTCTCATT

GTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGGCAGAGAGACTGGACAGCTCTGGA

GGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGCCTCTGGAGTGGGACACTGGCCCTG

GGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCGTTGGATCAGACCCTCCTGTGGGCAGGG

TTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATAAAAACCAACCCAAATCAA

FIGURE 520

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPGE
HAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGESQV
FKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYYHKL
RMSVEYSQSWGHFQNRVNLVGDIFRNDGSIMLQGVRESDGGNYTCSIHLGNLVFKKTIVLHVS
PEEPRTLVTPAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKTCGNKSSVNSTVLVKNT
KKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLRSDRNNS
LEKKSGGGMPKTQQAF

FIGURE 521

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATGG

TTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGGAA

ACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CAGGA

TGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTGGCCC

TGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGGATGGT

TGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGATGAGAA

TGAAAATCGCACAGGAACTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGGTAAAACA

ATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAAACTGGAGATA

TTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAGTAAGCAGTA

CTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGGAGTACATCAA

AGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAATGAGGTCTGGAA

GTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGAAGATGGAAAAGGAAA

TATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTGAGAACAAACATTA

TTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT**TAA**TGCAAAGAGGT

GGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAGATATGTATGAATGCATCAGTA

GCTGAAAAAAAAAAAAAA

FIGURE 522

MQDEDGYITLNIKTRKPALVSVGPASSSWWRVMALILLILCVGMVVGLVALGIWSVMQRNYLQ
DENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSCYGFFRHNLTWEES
KQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEFLEDG
KGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

FIGURE 523

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATG**GC

AAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAA

GAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTTTGGTATCCTGGCCCTAACTCTAAT

TGTCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAAGCCTATGACATGGA

GCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAG

AACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAA

CGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAAACTCAGATTAAAGT

GATTCCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACTTT

CTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAGCCTATTGAAAACCGAGATTTTCTTAA

AAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAAT

ATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGA

AAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCG

TCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTGAAAATGGAATAGA

ATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTA

TTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATACTGCTACCAAGGAGG

ACGAGTCATCTGTCGTGTCATCATGCCTTGTAACTGGTGGGTGGCCCGCATGCTGGGGAGGGT

C**TAA**TAGGAGGTTTGAGCTCAAATGCTTAAACTGCTGGCAACATATAATAAATGCATGCTATT

CAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCTCTCCAGAATTACTTGTAG

GTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACCAAAAAAAAAAAAAAAAAAA

FIGURE 524

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPKKAYD
MEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI
KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNSKILEICDNVTMYWINPT
LISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENG
IEFDPMLDERGYCCIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features:
Signal peptide:
amino acids 1-40

Transmembrane domain:
amino acids 25-47 (type II)

N-glycosylation sites.
amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.
amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.
amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.
amino acids 315-317

Cytochrome c family heme-binding site signature.
amino acids 9-14

FIGURE 525

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGC**ATG**AAGATCACTGGGGGTC

TCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCCAA

AAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCACAT

ACCTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACTTGTGTACCGAGA

GCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGC**TAA**ATTCTCCATGGAC

ATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTGAGTTTC

TTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTGACTGAAT

GGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCATTTTTTTTT

TAACACGTCAATAAAAAAATAATCTCCCAGA

FIGURE 526

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGNE
CHLCTESLKSNGRVQFLHDGSC

Important features:

Signal peptide:

amino acids 1-19

FIGURE 527

CGACG**ATG**CTACGCGCGCCCGGCTGCCTCCTCCGGACCTCCGTAGCGCCTGCCGCGGCCCTGG
CTGCGGCGCTGCTCTCGTCGCTTGCGCGCTGCTCTCTTCTAGAGCCGAGGGACCCGGTGGCCT
CGTCGCTCAGCCCCTATTTCGGCACCAAGACTCGCTACGAGGATGTCAACCCCGTGCTATTGT
CGGGCCCCGAGGCTCCGTGGCGGGACCCTGAGCTGCTGGAGGGGACCTGCACCCCGGTGCAGC
TGGTCGCCCTCATTCGCCACGGCACCCGCTACCCCACGGTCAAACAGATCCGCAAGCTGAGGC
AGCTGCACGGGTTGCTGCAGGCCCGCGGGTCCAGGGATGGCGGGGCTAGTAGTACCGGCAGCC
GCGACCTGGGTGCAGCGCTGGCCGACTGGCCTTTGTGGTACGCGGACTGGATGGACGGGCAGC
TAGTAGAGAAGGGACGGCAGGATATGCGACAGCTGGCGCTGCGTCTGGCCTCGCTCTTCCCGG
CCCTTTTCAGCCGTGAGAACTACGGCCGCCTGCGGCTCATCACCAGTTCCAAGCACCGCTGCA
TGGATAGCAGCGCCGCCTTCCTGCAGGGGCTGTGGCAGCACTACCACCCTGGCTTGCCGCCGC
CGGACGTCGCAGATATGGAGTTTGGACCTCCAACAGTTAATGATAAACTAATGAGATTTTTTG
ATCACTGTGAGAAGTTTTTAACTGAAGTAGAAAAAAATGCTACAGCTCTTTATCACGTGGAAG
CCTTCAAAACTGGACCAGAAATGCAGAACATTTTAAAAAAGTTGCAGCTACTTTGCAAGTGC
CAGTAAATGATTTAAATGCAGATTTAATTCAAGTAGCCTTTTTCACCTGTTCATTTGACCTGG
CAATTAAAGGTGTTAAATCTCCTTGGTGTGATGTTTTTGACATAGATGATGCAAAGGTATTAG
AATATTTAAATGATCTGAAACAATATTGGAAAAGAGGATATGGGTATACTATTAACAGTCGAT
CCAGCTGCACCTTGTTTCAGGATATCTTTCAGCACTTGGACAAAGCAGTTGAACAGAAACAAA
GGTCTCAGCCAATTTCTTCTCCAGTCATCCTCCAGTTTGGTCATGCAGAGACTCTTCTTCCAC
TGCTTTCTCATGGGCTACTTCAAAGACAAGGAACCCCTAACAGCGTACAATTACAAAAAAC
AAATGCATCGGAAGTTCCGAAGTGGTCTCATTGTACCTTATGCCTCGAACCTGATATTTGTGC
TTTACCACTGTGAAAATGCTAAGACTCCTAAAGAACAATTCCGAGTGCAGATGTTATTAAATG
AAAAGGTGTTACCTTTGGCTTACTCACAAGAAACTGTTTCATTTTATGAAGATCTGAAGAACC
ACTACAAGGACATCCTTCAGAGTTGTCAAACCAGTGAAGAATGTGAATTAGCAAGGGCTAACA
GTACATCTGATGAACTA**TGA**GTAACTGAAGAACATTTTTAATTCTTTAGGAATCTGCAATGAG
TGATTACATGCTTGTAATAGGTAGGCAATTCCTTGATTACAGGAAGCTTTTATATTACTTGAG
TATTTCTGTCTTTTCACAGAAAAACATTGGGTTTCTCTCTGGGTTTGGACATGAAATGTAAGA
AAAGATTTTTCACTGGAGCAGCTCTCTTAAGGAGAAACAAATCTATTTAGAGAAACAGCTGGC
CCTGCAAATGTTTACAGAAATGAAATTCTTCCTACTTATATAAGAAATCTCACACTGAGATAG
AATTGTGATTTCATAATAACACTTGAAAAGTGCTGGAGTAACAAAATATCTCAGTTGGACCAT
CCTTAACTTGATTGAACTGTCTAGGAACTTTACAGATTGTTCTGCAGTTCTCTCTTCTTTTCC
TCAGGTAGGACAGCTCTAGCATTTTCTTAATCAGGAATATTGTGGTAAGCTGGGAGTATCACT
CTGGAAGAAAGTAACATCTCCAGATGAGAATTTGAAACAAGAAACAGAGTGTTGTAAAAGGAC
ACCTTCACTGAAGCAAGTCGGAAAGTACAATGAAAATAAATATTTTTGGTATTTATTTATGAA
ATATTTGAACATTTTTTCAATAATTCCTTTTTACTTCTAGGAAGTCTCAAAAGACCATCTTAA
ATTATTATATGTTTGGACAATTAGCAACAAGTCAGATAGTTAGAATCGAAGTTTTTCAAATCC
ATTGCTTAGCTAACTTTTTCATTCTGTCACTTGGCTTCGATTTTTATATTTTCCTATTATATG
AAATGTATCTTTTGGTTGTTTGATTTTTCTTTCTTTCTTTGTAAATAGTTCTGAGTTCTGTCA
AATGCCGTGAAAGTATTTGCTATAATAAAGAAATTCTTGTGACTTTAAAAAAAAA

FIGURE 528

MLRAPGCLLRTSVAPAAALAAALLSSLARCSLLEPRDPVASSLSPYFGTKTRYEDVNPVLLSG
PEAPWRDPELLEGTCTPVQLVALIRHGTRYPTVKQIRKLRQLHGLLQARGSRDGGASSTGSRD
LGAALADWPLWYADWMDGQLVEKGRQDMRQLALRLASLFPALFSRENYGRLRLITSSKHRCMD
SSAAFLQGLWQHYHPGLPPPDVADMEFGPPTVNDKLMRFFDHCEKFLTEVEKNATALYHVEAF
KTGPEMQNILKKVAATLQVPVNDLNADLIQVAFFTCSFDLAIKGVKSPWCDVFDIDDAKVLEY
LNDLKQYWKRGYGYTINSRSSCTLFQDIFQHLDKAVEQKQRSQPISSPVILQFGHAETLLPLL
SLMGYFKDKEPLTAYNYKKQMHRKFRSGLIVPYASNLIFVLYHCENAKTPKEQFRVQMLLNEK
VLPLAYSQETVSFYEDLKNHYKDILQSCQTSEECELARANSTSDEL

Important features:

Signal sequence

amino acids 1-30

N-glycosylation sites.

amino acids 242-246, 481-485

N-myristoylation sites.

amino acids 107-113, 113-119, 117-123, 118-124, 128-134

Endoplasmic reticulum targeting sequence.

amino acids 484-489

FIGURE 529

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGGTGG
CGGAGAGATCAGAAGCCTCTTCCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCTCAGGGA
CGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACG**ATG**GCAGCGTCCGCCGGAGCCGGGGCG
GTGATTGCAGCCCCAGACAGCCGGCGCTGGCTGTGGTCGGTGCTGGCGGCGGCGCTTGGGCTC
TTGACAGCTGGAGTATCAGCCTTGGAAGTATATACGCCAAAAGAAATCTTCGTGGCAAATGGT
ACACAAGGGAAGCTGACCTGCAAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACCTCAGTC
TCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTCGTTTTTCCACTACTCCCAAGGG
CAAGTGTACCTTGGGAATTATCCACCATTTAAAGACAGAATCAGCTGGGCTGGAGACCTTGAC
AAGAAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACACAATGGCACCTATATCTGT
GATGTCAAAAACCCTCCTGACATCGTTGTCCAGCCTGGACACATTAGGCTCTATGTCGTAGAA
AAAGAGAATTTGCCTGTGTTTCCAGTTTGGGTAGTGGTGGGCATAGTTACTGCTGTGGTCCTA
GGTCTCACTCTGCTCATCAGCATGATTCTGGCTGTCCTCTATAGAAGGAAAAACTCTAAACGG
GATTACACTGGCTGCAGTACATCAGAGAGTTTGTCACCAGTTAAGCAGGCTCCTCGGAAGTCC
CCCTCCGACACTGAGGGTCTTGTAAAGAGTCTGCCTTCTGGATCTCACCAGGGCCCAGTCATA
TATGCACAGTTAGACCACTCCGGCGGACATCACAGTGACAAGATTAACAAGTCAGAGTCTGTG
GTGTATGCGGATATCCGAAAGAAT**TAA**GAGAATACCTAGAACATATCCTCAGCAAGAAACAAA
ACCAAACTGGACTCTCGTGCAGAAAATGTAGCCCATTACCACATGTAGCCTTGGAGACCCAGG
CAAGGACAAGTACACGTGTACTCACAGAGGGAGAGAAAGATGTGTACAAAGGATATGTATAAA
TATTCTATTTAGTCATCCTGATATGAGGAGCCAGTGTTGCATGATGAAAAGATGGTATGATTC
TACATATGTACCCATTGTCTTGCTGTTTTTGTACTTTCTTTTCAGGTCATTTACAATTGGGAG
ATTTCAGAAACATTCCTTTCACCATCATTTAGAAATGGTTTGCCTTAATGGAGACAATAGCAG
ATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAATCTAAGGGCTTAAGACTGATTAGTCTTA
GCATTTACTGTAGTTGGAGGATGGAGATGCTATGATGGAAGCATACCCAGGGTGGCCTTTAGC
ACAGTATCAGTACCATTTATTTGTCTGCCGCTTTTAAAAAATACCCATTGGCTATGCCACTTG
AAAACAATTTGAGAAGTTTTTTTGAAGTTTTTCTCACTAAAATATGGGGCAATTGTTAGCCTT
ACATGTTGTGTAGACTTACTTTAAGTTTGCACCCTTGAAATGTGTCATATCAATTTCTGGATT
CATAATAGCAAGATTAGCAAAGGATAAATGCCGAAGGTCACTTCATTCTGGACACAGTTGGAT
CAATACTGATTAAGTAGAAAATCCAAGCTTTGCTTGAGAACTTTTGTAACGTGGAGAGTAAAA
AGTATCGGTTTTA

FIGURE 530

MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKST
STTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDKKDASINIENMQ
FIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLTLLISMILAV
LYRRKNSKRDYTGCSTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSGGHHS
DKINKSESVVYADIRKN

Important features:

Signal peptide:

amino acids 1-37

Transmembrane domain:

amino acids 161-183

FIGURE 531

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCTCGA

GGCTGGTGGGAAGAAGCCGAG**ATG**GCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGCTC

TTGCTGATGGCGGTAGCAGCGCCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACTGGT

GCGCGAGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCTGTGGCACGGTGGGGCTGCTGCTG

GAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCTGGAAC

CAGCAGGATGGTACCTTGTCCCTGTCACAGCGGCAGCTCAGCGAGGAGGAGCGGGGCCGACTC

CGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGGGGCCCTG

GATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGGTGGAGTCG

CACCTGTCGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGCGTGTCGGTG

GTGACGCACCCCGGGGGCTGCCGGGGCCATGAGGTGGAGGACGTGGACCTGGAGCTGTTCAAC

ACCTCGGTGCAGCTGCAGCCGCCCACCACAGCCCCAGGCCCTGAGACGGCGGCCTTCATTGAG

CGCCTGGAGATGGAACAGGCCCAGAAGGCCAAGAACCCCCAGGAGCAGAAGTCCTTCTTCGCC

AAATACTGGATGTACATCATTCCCGTCGTCCTGTTCCTCATGATGTCAGGAGCGCCAGACACC

GGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCCTTTGCTGTGTGCCA

CCCTCCCTG**TAA**GTCTATTTAAAAACATCGACGATACATTGAAATGTGTGAACGTTTTGAAAA

GCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGGTCCTGATGTACAAGCT

TGATTGAAATTCACTGCTCACTTGATACGTTATTCAGAAACCCAAGGAATGGCTGTCCCCATC

CTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTATTAAACTGTCCCCCAGATC

GACACGCAAAAAAAAA

FIGURE 532

MAAASAGATRLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEID
DSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVRIPRRPGALDGLEAGG
YVSSFVPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELFNTSVQLQP
PTTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFAKYWMYIIPVVLFLMMSGAPDTGGQGGGG
GGGGGGGSGLCCVPPSL

Important features:

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 226-243

FIGURE 533

TCTGCCTCCACTGCTCTGTGCTGGGATC**ATG**GAACTTGCACTGCTGTGTGGGCTGGTGGTGAT

GGCTGGTGTGATTCCAATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTGAC

TGGGAAAATGCCCATCCTCTCCTACTGGCCCTACGGCTGTCACTGCGGACTAGGTGGCAGAGG

CCAACCCAAAGATGCCACGGACTGGTGCTGCCAGACCCATGACTGCTGCTATGACCACCTGAA

GACCCAGGGGTGCGGCATCTACAAGGACAACAACAAAGCAGCATACATTGTATGGATTTATC

TCAACGCTATTGTTTAATGGCTGTGTTTAATGTGATCTATCTGGAAAATGAGGACTCCGAA**TA**

AAAAGCTATTACTAWTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 534

MELALLCGLVVMAGVIPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDWC
CQTHDCCYDHLKTQGCGIYKDNNKSSIHCMDLSQRYCLMAVFNVIYLENEDSE

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 1-24

N-glycosylation site.

amino acids 86-89

N-myristoylation sites.

amino acids 20-25, 45-50

Phospholipase A2 histidine active site.

amino acids 63-70

FIGURE 535

```
GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTG
AACTGGGTGCTCATCACGGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCC
CCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTT
TTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAACTGTAAAGATGCAAAAACGTAATATCCAT
GAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTT
GTTCTTGGAGTGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAG
GGGTCCAATTTTTCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGG
CTGTCATGCAACTGGCCCCTAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAACAATACAA
AGGATGGGTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCC
ACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAA
GGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGT
TGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAACTTAAGTATAATCAATTTAAAGGG
CTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATATTGACGAAAATGCT
TTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTT
AACAATACCTTCAGACCTGTGACAAATTTACGGAACTTGGATCTGTCCTATAATCAGCTGCAT
TCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTTACATTTACGGTCTAAC
TCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACCTGGAACTTTTGGACCTG
GGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGCATGATCAGACTCAAAGAA
CTTCACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCTTTTTCCAAGGTTGGTCAGC
CTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTCATAGGACAGACCATGTCCTGGACC
TGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATCGAAGCTTTCAGTGGACCCAGT
GTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGATTCCAACAAGCTCACATTTATT
GGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAGTCTTGCTGGGAATATATGG
GAATGCAGCAGAAATATTTGCTCCCTTGTAAACTGGCTGAAAAGTTTTAAAGGTCTAAGGGAG
AATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTGATCGATGCAGTGAAG
AACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAGGGCTCTCCCAAAG
CCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCCTTTGCCCCCGACG
GTGGGAGCCACAGAGCCCGGCCCAGAGACCGATGCTGACGCCGAGCACATCTCTTTCCATAAA
ATCATCGCGGGCAGCGTGGCGCTTTTCCTGTCCGTGCTCGTCATCCTGCTGGTTATCTACGTG
TCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAAGGCAC
AGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTAGAT
TATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTAT
AACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAGAGCTCTTAAAAGC
TGGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCCTC
CCCTTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTTGTCCGTTTTAGTGCATTCATA
ATACTGGTCATTTTCCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGT
GAAGCTTGAACTCCGGTTTAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAAT
GTCGCATTTGTTTTAAGATAAAACTTCTTTCATAGGTAAAAAAAAAA
```

FIGURE 536

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGC
LGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLN
NTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELLDLG
YNRIRSLARNVFAGMIRLKELHLEHNQFSKLNLALFPRLVSLQNLYLQWNKISVIGQTMSWTW
SSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDSNKLTFIGQEILDSWISLNDISLAGNIWE
CSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKP
TFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVS
WKRYPASMKQLQQRSLMRRHRKKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCTYN
KSGSRECEV

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 537

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCTCAGCAACCCTCGAC**ATG**GCGCTGAGGCGGCCACCGCGAC
TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCCTGCTGCTGCTTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC
TCAAATCCAGCAATCGAACCCCAGTGGTACAGGAATTTGAAAGTGTGGAACTGTCTTGCATCATTACGGATTCGC
AGACAAGTGACCCCAGGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTGACAACAAAA
TTCAGGGAGACTTGGCGGGTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG
ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC
ACTGCCAGGAGAGTGAGGGCCACCCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCCACGGATT
CCAGAGCCAATCCCAGATTTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCACTGCTG
TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG
AGATGGAAGTCTATGACCTGAACATTGGCGGAATTATTGGGGGGGTTCTGGTTGTCCTTGCTGTACTGGCCCTGA
TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA
ACCCAGGGAAACCAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTG
TGATC**TGA**GACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTCAGAAGCTTTTCGTTTTGGCCAAAGTTGACCA
CTACTCTTCTTACTCTAACAAGCCACATGAATAGAAGAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTCTGGCCTGATTCCCGCATGAGTATTAGG
GTGATCTTAAAGAGTTTGCTCACGTAAACGCCCGTGCTGGGCCCTGTGAAGCCAGCATGTTCACCACTGGTCGTT
CAGCAGCCACGACAGCACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGGCGGGAACCCA
GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC
TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG
GTAAATTGGTTGCTGGAAGAGGGATCTTGCCTGAGGAACCCTGCTTGTCCAACAGGGTGTCAGGATTTAAGGAAA
ACCTTCGTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTTTATTTTATAAAATTT
TACATCTAAATTTTTGCTAAGGATGTATTTTGATTATTGAAAGAAAATTTCTATTTAAACTGTAAATATATTGT
CATACAATGTTAAATAACCTATTTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTTAAAT
TGGAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT
CACACAAGTTTTAGCCTTTTTCACAAGGGAACTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT
TAAAAATTCCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT
GCCTCTTCCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG
CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT
GCCATCCTTGGGCCCTGGCAGTGGCTGTGTCCCAGTGAGCTTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC
TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT
TTTTGGTTATGGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTTAAGTTTGTTTAATTATTTGTT
AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTAAAAGAAAATGGAT
CCCACTGTTCCTCTTTGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTCGCATTTCAAAACAAACCATGAT
GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTCAGGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG
TGAAACGCCTGAATCAAAAGCAGTTTTCTAATTTTTGACTTTAAATTTTTCATCCGCCGGAGACACTGCTCCCATT
TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAGAGCAGGTGTTCTCAGCCTCACATGCCCT
GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCTGGA
GAATGGCTCTCACTACTCACCTTGTCTTTCAGCTTCCAGTGTCTTGGGTTTTTTATACTTTGACAGCTTTTTTTT
AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTTATGTGAAACACTTTGCCGCAGGCCGCCTGGCAGAGGCA
GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCATGGCATCCTGGATGCTTAGCATGCAAGTTC
CCTCCATCATTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTCACGCTCCAGCCTCCT
TCTTGGTTGTCATAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA
TGGGAACCAGGTCTGAAAAAGTAGAGAGAAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA
CGGAAAAGGAATACTCGTGTATTTTAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTTGTTTAAC
CTCATTTATAAAGCTTCAAAAAAACCCA

FIGURE 538

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTSDP

RIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLKIWNVTRRDSALYRCEVVARNDRKEI

DEIVIELTVQVKPVTPVCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRANPR

FRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVYDLNIGGIIGGVLVV

LAVLALITLGICCAYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Important features:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267

FIGURE 539

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAGAGGCCGGGGAAGAGAAGCAAAGC
GCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCCTAA
CTTCAGTCCCCCAAACGCGCACCCTCGAAGTCTTGAACTCCAGCCCCGCACATCCACGCGCGG
CACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCGGGCAGCTGG
GCTCGGGCGGCGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGG
CTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCG**A**
TGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCC
GCCGCGTGGTCAGCGGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGG
CCTACTTCCATGAACTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTG
AGGGAGGAGTCCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGC
AAAACCTGACAAAACCCGGGACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGA
ATGGAGATGGGCAAACATCTGGTGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATT
CCCAGTACCGAAACTGGTACACAGATGAACCTTCCTGCGGAAGTGAAAGTGTGTTGTGATGT
ATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACA
GGTGTAACATGAAGCACAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCTG
TAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAG
CAGGTATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATAC
TGGTTGCTTTTGGAACCTGTTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAAACTA
GTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAGAAAGTGGCATGGAAGTA**T**
AATAACTCATTGACTTGGTTCCAGAATTTTGTAATTCTGGATCTGTATAAGGAATGGCATCAG
AACAATAGCTTGGAATGGCTTGAAATCACAAAGGATCTGCAAGATGAACTGTAAGCTCCCCCT
TGAGGCAAATATTAAAGTAATTTTTATATGTCTATTATTTCATTTAAAGAATATGCTGTGCTA
ATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAAACTTCAAACTTCAAGCAAA
TGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTATGTGTGTTAGAAGCAAT
TCCTTTTATTTCTTTCACCTTTCATAAGTTGTTATCTAGTCAATGTAATGTATATTGTATTGA
AATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTTTGATAAAAATGAACTGTTCTA
ATATTTATTTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAGAAACTTATTAC
TGTTGTCAACTGAATTCACACACACAAATATAGTACCATAGAAAAAGTTTGTTTTCTCGAA
ATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGAAATAAGA
AGCTATTTCATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGATTGTCT
AATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTTAGCTTAAAATTAAACAGATTTTG
TAATAATGTAACTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAGT
GACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTG
AGGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCGGGGTTT
GGGATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGC
CTCTGACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAG
TGTGGAGACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACTACGAAATCGTGTGAAA
ATGGGTTGGAACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATG
GTGGCTCCTTTCTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTTCTTCTCAAGAGA
AAGTTGTAACTCTCTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAAATAAGAGTTCTTG
TTTCTGGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 540

MSRVVSLLLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACES
EGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSN
SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAP
VEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKT
SPNQSTLWISKSTRKESGMEV

Important features:
Signal peptide:
amino acids 1-21

Transmembrane domain:
amino acids 214-235

N-glycosylation sites.
amino acids 86-89 and 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 266-269

N-myristoylation sites.
amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145
and 212-217

FIGURE 541

GGAGA**ATG**GAGAGAGCAGTGAGAGTGGAGTCCGGGGTCCTGGTCGGGGTGGTCTGTCTGCTCCTGGCATGCCCTG
CCACAGCCACTGGGCCCGAAGTTGCTCAGCCTGAAGTAGACACCACCCTGGGTCGTGTGCGAGGCCGGCAGGTGG
GCGTGAAGGGCACAGACCGCCTTGTGAATGTCTTTCTGGGCATTCCATTTGCCCAGCCGCCACTGGGCCCTGACC
GGTTCTCAGCCCCACACCCAGCACAGCCCTGGGAGGGTGTGCGGGATGCCAGCACTGCGCCCCCAATGTGCCTAC
AAGACGTGGAGAGCATGAACAGCAGCAGATTTGTCCTCAACGGAAAACAGCAGATCTTCTCCGTTTCAGAGGACT
GCCTGGTCCTCAACGTCTATAGCCCAGCTGAGGTCCCCGCAGGGTCCGGTAGGCCGGTCATGGTATGGGTCCATG
GAGGCGCTCTGATAACTGGCGCTGCCACCTCCTACGATGGATCAGCTCTGGCTGCCTATGGGGATGTGGTCGTGG
TTACAGTCCAGTACCGCCTTGGGGTCCTTGGCTTCTTCAGCACTGGAGATGAGCATGCACCTGGCAACCAGGGCT
TCCTAGATGTGGTAGCTGCTTTGCGCTGGGTGCAAGAAAACATCGCCCCCTTCGGGGGTGACCTCAACTGTGTCA
CTGTCTTTGGTGGATCTGCCGGTGGGAGCATCATCTCTGGCCTGGTCCTGTCCCCAGTGGCTGCAGGGCTGTTCC
ACAGAGCCATCACACAGAGTGGGGTCATCACCACCCCAGGGATCATCGACTCTCACCCTTGGCCCCTAGCTCAGA
AAATCGCAAACACCTTGGCCTGCAGCTCCAGCTCCCCGGCTGAGATGGTGCAGTGCCTTCAGCAGAAAGAAGGAG
AAGAGCTGGTCCTTAGCAAGAAGCTGAAAAATACTATCTATCCTCTCACCGTTGATGGCACTGTCTTCCCCAAAA
GCCCCAAGGAACTCCTGAAGGAGAAGCCCTTCCACTCTGTGCCCTTCCTCATGGGTGTCAACAACCATGAGTTCA
GCTGGCTCATCCCCAGGGGCTGGGGTCTCCTGGATACAATGGAGCAGATGAGCCGGGAGGACATGCTGGCCATCT
CAACACCCGTCTTGACCAGTCTGGATGTGCCCCCTGAGATGATGCCCACCGTCATAGATGAATACCTAGGAAGCA
ACTCGGACGCACAAGCCAAATGCCAGGCGTTCCAGGAATTCATGGGTGACGTATTCATCAATGTTCCCACCGTCA
GTTTTTCAAGATACCTTCGAGATTCTGGAAGCCCTGTCTTTTTTCTATGAGTTCCAGCATCGACCCAGTTCTTTTG
CGAAGATCAAACCTGCCTGGGTGAAGGCTGATCATGGGGCCGAGGGTGCTTTTGTGTTCGGAGGTCCCTTCCTCA
TGGACGAGAGCTCCCGCCTGGCCTTTCCAGAGGCCACAGAGGAGGAGAAGCAGCTAAGCCTCACCATGATGGCCC
AGTGGACCCACTTTGCCCGGACAGGGGACCCCAATAGCAAGGCTCTGCCTCCTTGGCCCCAATTCAACCAGGCGG
AACAATATCTGGAGATCAACCCAGTGCCACGGGCCGGACAGAAGTTCAGGGAGGCCTGGATGCAGTTCTGGTCAG
AGACGCTCCCCAGCAAGATACAACAGTGGCACCAGAAGCAGAAGAACAGGAAGGCCCAGGAGGACCTC**TGA**GGCC
AGGCCTGAACCTTCTTGGCTGGGGCAAACCACTCTTCAAGTGGTGGCAGAGTCCCAGCACGGCAGCCCGCCTCTC
CCCCTGCTGAGACTTTAATCTCCACCAGCCCTTAAAGTGTCGGCCGCTCTGTGACTGGAGTTATGCTCTTTTGAA
ATGTCACAAGGCCGCCTCCCACCTCTGGGGCATTGTACAAGTTCTTCCCTCTCCCTGAAGTGCCTTTCCTGCTTT
CTTCGTGGTAGGTTCTAGCACATTCCTCTAGCTTCCTGGAGGACTCACTCCCCAGGAAGCCTTCCCTGCCTTCTC
TGGGCTGTGCGGCCCCGAGTCTGCGTCCATTAGAGCACAGTCCACCCGAGGCTAGCACCGTGTCTGTGTCTGTCT
CCCCCTCAGAGGAGCTCTCTCAAAATGGGGATTAGCCTAACCCCACTCTGTCACCCACACCAGGATCGGGTGGGA
CCTGGAGCTAGGGGGTGTTTGCTGAGTGAGTGAGTGAAACACAGAATATGGGAATGGCAGCTGCTGAACTTGAAC
CCAGAGCCTTCAGGTGCCAAAGCCATACTCAGGCCCCCACCGACATTGTCCACCCTGGCCAGAAGGGTGCATGCC
AATGGCAGAGACCTGGGATGGGAGAAGTCCTGGGGCGCCAGGGGATCCAGCCTAGAGCAGACCTTAGCCCCTGAC
TAAGGCCTCAGACTAGGGCGGGAGGGGTCTCCTCCTCTCTGCTGCCCAGTCCTGGCCCCTGCACAAGACAACAGA
ATCCATCAGGGCCATGAGTGTCACCCAGACCTGACCCTCACCAATTCCAGCCCCTGACCCTCAGGACGCTGGATG
CCAGCTCCCAGCCCCAGTGCCGGGTCCTCCCTCCCTTCCTGGCTTGGGGGAGACCAGTTTCTGGGGAGCTTCCAAG
AGCACCCACCAAGACACAGCAGGACAGGCCAGGGGAGGGCATCTGGACCAGGGCATCCGTCGGGCTATTGTCACA
GAGAAAAGAAGAGACCCACCCACTCGGGCTGCAAAAGGTGAAAAGCACCAAGAGGTTTTCAGATGGAAGTGAGAG
GTGACAGTGTGCTGGCAGCCCTCACAGCCCTCGCTTGCTCTCCCTGCCGCCTCTGCCTGGGCTCCCACTTTGGCA
GCACTTGAGGAGCCCTTCAACCCGCCGCTGCACTGTAGGAGCCCCTTTCTGGGCTGGCCAAGGCCGGAGCCAGCT
CCCTCAGCTTGCGGGGAGGTGCGGAGGGAGAGGGGCGGGCAGGAACCGGGGCTGCGCGCAGCGCTTGCGGGCCAG
AGTGAGTTCCGGGTGGGCGTGGGCTCGGCGGGGCCCACTCAGAGCAGCTGGCCGGCCCCAGGCAGTGAGGGCCT
TAGCACCTGGGCCAGCAGCTGCTGTGCTCGATTTCTCGCTGGGCCTTAGCTGCCTCCCCGCGGGGCAGGGCTCGG
GACCTGCAGCCCTCCATGCCTGACCCTCCCCCCACCCCCGTGGGCTCCTGTGCGGCCGGAGCCTCCCCAAGGAG
CGCCGCCCCCTGCTCCACAGCGCCCAGTCCCATCGACCACCCAAGGGCTGAGGAGTGCGGGTGCACAGCGCGGGA
CTGGCAGGCAGCTCCACCTGCTGCCCCAGTGCTGGATCCACTGGGTGAAGCCAGCTGGGCTCCTGAGTCTGGTGG
GGACTTGGAGAACCTTTATGTCTAGCTAAGGGATTGTAAATACACCGATGGGCACTCTGTATCTAGCTCAAGGTT
TGTAAACACACCAATCAGCACCCTGTGTCTAGCTCAGTGTTTGTGAATGCACCAATCCACACTCTGTATCTGGCT
ACTCTGGTGGGGACTTGGAGAACCTTTGTGTCCACACTCTGTATCTAGCTAATCTAGTGGGGATGTGGAGAACCT
TTGTGTCTAGCTCAGGGATCGTAAACGCACCAATCAGCACCCTGTCAAAACAGACCACTTGACTCTCTGTAAAAT
GGACCAATCAGCAGGATGTGGGTGGGGCGAGACAAGAGAATAAAAGCAGGCTGCCTGAGCCAGCAGTGACAACCC.
CCCTCGGGTCCCCTCCCACGCCGTGGAAGCTTTGTTCTTTCGCTCTTTGCAATAAATCTTGCTACTGCCCAAAA

FIGURE 542

MERAVRVESGVLVGVVCLLLACPATATGPEVAQPEVDTTLGRVRGRQVGVKGTDRLVNVFLGI
PFAQPPLGPDRFSAPHPAQPWEGVRDASTAPPMCLQDVESMNSSRFVLNGKQQIFSVSEDCLV
LNVYSPAEVPAGSGRPVMVWVHGGALITGAATSYDGSALAAYGDVVVVTVQYRLGVLGFFSTG
DEHAPGNQGFLDVVAALRWVQENIAPFGGDLNCVTVFGGSAGGSIISGLVLSPVAAGLFHRAI
TQSGVITTPGIIDSHPWPLAQKIANTLACSSSSPAEMVQCLQQKEGEELVLSKKLKNTIYPLT
VDGTVFPKSPKELLKEKPFHSVPFLMGVNNHEFSWLIPRGWGLLDTMEQMSREDMLAISTPVL
TSLDVPPEMMPTVIDEYLGSNSDAQAKCQAFQEFMGDVFINVPTVSFSRYLRDSGSPVFFYEF
QHRPSSFAKIKPAWVKADHGAEGAFVFGGPFLMDESSRLAFPEATEEEKQLSLTMMAQWTHFA
RTGDPNSKALPPWPQFNQAEQYLEINPVPRAGQKFREAWMQFWSETLPSKIQQWHQKQKNRKA
QEDL

Important features:

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 226-245

N-glycosylation site.

amino acids 105-109

N-myristoylation sites.

amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161, 162-168, 217-223, 227-233, 228-234, 232-238, 262-268, 357-363, 461-467

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23

Carboxylesterases type-B serine active site.

amino acids 216-232

FIGURE 543

TGTCGCCTGGCCCTCGCC**ATG**CAGACCCCGCGAGCGTCCCCTCCCCGCCCGGCCCTCCTGCTTCTGCTGCTGCTA
CTGGGGGGCGCCCACGGCCTCTTTCCTGAGGAGCCGCCGCCGCTTAGCGTGGCCCCCAGGGACTACCTGAACCAC
TATCCCGTGTTTGTGGGCAGCGGGCCCGGACGCCTGACCCCCGCAGAAGGTGCTGACGACCTCAACATCCAGCGA
GTCCTGCGGGTCAACAGGACGCTGTTCATTGGGGACAGGGACAACCTCTACCGCGTAGAGCTGGAGCCCCCCACG
TCCACGGAGCTGCGGTACCAGAGGAAGCTGACCTGGAGATCTAACCCCAGCGACATAAACGTGTGTCGGATGAAG
GGCAAACAGGAGGGCGAGTGTCGAAACTTCGTAAAGGTGCTGCTCCTTCGGGACGAGTCCACGCTCTTTGTGTGC
GGTTCCAACGCCTTCAACCCGGTGTGCGCCAACTACAGCATAGACACCCTGCAGCCCGTCGGAGACAACATCAGC
GGTATGGCCCGCTGCCCGTACGACCCCAAGCACGCCAATGTTGCCCTCTTCTCTGACGGGATGCTCTTCACAGCT
ACTGTTACCGACTTCCTAGCCATTGATGCTGTCATCTACCGCAGCCTCGGGGACAGGCCCACCCTGCGCACCGTG
AAACATGACTCCAAGTGGTTCAAAGAGCCTTACTTTGTCCATGCGGTGGAGTGGGGCAGCCATGTCTACTTCTTC
TTCCGGGAGATTGCGATGGAGTTTAACTACCTGGAGAAGGTGGTGGTGTCCCGCGTGGCCCGAGTGTGCAAGAAC
GACGTGGGAGGCTCCCCCCGCGTGCTGGAGAAGCAGTGGACGTCCTTCCTGAAGGCGCGGCTCAACTGCTCTGTA
CCCGGAGACTCCCATTTCTACTTCAACGTGCTGCAGGCTGTCACGGGCGTGGTCAGCCTCGGGGGCCGGCCCGTG
GTCCTGGCCGTTTTTTTCCACGCCCAGCAACAGCATCCCTGGCTCGGCTGTCTGCGCCTTTGACCTGACACAGGTG
GCAGCTGTGTTTGAAGGCCGCTTCCGAGAGCAGAAGTCCCCCGAGTCCATCTGGACGCCGGTGCCGGAGGATCAG
GTGCCTCGACCCCGGCCCGGGTGCTGCGCAGCCCCCGGGATGCAGTACAATGCCTCCAGCGCCTTGCCGGATGAC
ATCCTCAACTTTGTCAAGACCCACCCTCTGATGGACGAGGCGGTGCCCTCGCTGGGCCATGCGCCCTGGATCCTG
CGGACCCTGATGAGGCACCAGCTGACTCGAGTGGCTGTGTGGACGTGGGAGCCGGCCCCTGGGGCAACCAGACCGTT
GTCTTCCTGGGTTCTGAGGCGGGGACGGTCCTCAAGTTCCTCGTCCGGCCCAATGCCAGCACCTCAGGGACGTCT
GGGCTCAGTGTCTTCCTGGAGGAGTTTGAGACCTACCGGCCGGACAGGTGTGGACGGCCCGGCGGTGGCGAGACA
GGGCAGCGGCTGCTGAGCTTGGAGCTGGACGCAGCTTCGGGGGGCGTGCTGGCTGGCTGCCTTCCCCCGCTGCGTGGTC
CGAGTGCCTGTGGCTCGCTGCCAGCAGTACTCGGGGTGTATGAAGAACTGTATCGGCAGTCAGGACCCCTACTGC
GGGTGGGCCCCCGACGGCTCCTGCATCTTCCTCAGCCCGGGCACCAGAGCCGCCTTTGAGCAGGACGTGTCCGGG
GCCAGCACCTCAGGCTTAGGGGACTGCACAGGACTCCTGCGGGCCAGCCTCTCCGAGGACCGCGGGGCTGGTG
TCGGTGAACCTGCTGGTAACGTCGTCGGTGGCGGCCTTCGTGGTGGGAGCCGTGGTGTCCGGCTTCAGCGTGGGC
TGGTTCGTGGGCCTCCGTGAGCGGCGGGAGCTGGCCCGGCGCAAGGACAAGGAGGCCATCCTGGCGCACGGGGCG
GGCGAGGCGGTGCTGAGCGTCAGCCGCCTGGGCGAGCGCAGGGCGCAGGGTCCCGGGGGCCGGGGCGGAGGCGGT
GGCGGTGGCGCCGGGGTTCCCCCGGAGGCCCTGCTGGCGCCCCTGATGCAGAACGGCTGGGCCAAGGCCACGCTG
CTGCAGGGCGGGCCCCACGACCTGGACTCGGGGCTGCTGCCCACGCCCGAGCAGACGCCGCTGCCGCAGAAGCGC
CTGCCCACTCCGCACCCGCACCCCCACGCCCTGGGCCCCCGCGCCTGGGACCACGGCCACCCCCTGCTCCCGGCC
TCCGCTTCATCCTCCCTCCTGCTGCTGGCGCCCGCCCGGGCCCCCGAGCAGCCCCCCGCGCCTGGGGAGCCGACC
CCCGACGGCCGCCTCTATGCTGCCCGGCCCGGCCGCGCCTCCCACGGCGACTTCCCGCTCACCCCCCACGCCAGC
CCGGACCGCCGGCGGGTGGTGTCCGCGCCCACGGGCCCCTTGGACCCAGCCTCAGCCGCCGATGGCCTCCCGCGG
CCCTGGAGCCCGCCCCCGACGGGCAGCCTGAGGAGGCCACTGGGCCCCCACGCCCCTCCGGCCGCCACCCTGCGC
CGCACCCACACGTTCAACAGCGGCGAGGCCCGGCCTGGGGACCGCCACCGCGGCTGCCACGCCCGGCCGGGCACA
GACTTGGCCCACCTCCTCCCCTATGGGGGCGGACAGGACTGCGCCCCCCGTGCCC**TAG**GCCGGGGGCCCCCCG
ATGCCTTGGCAGTGCCAGCCACGGGAACCAGGAGCGAGAGACAGGTGCCAGAACGCCGGGGCCCGGGGCAACTCCG
AGTGGGTGCTCAAGTCCCCCCCGCGACCCACCCGCGGAGTGGGGGGCCCCCTCCGCCACAAGGAAGCACAACCAG
CTCGCCCTCCCCCTACCCGGGGCCGCAGGACGCTGAGACGGTTTGGGGGTGGGTGGGCGGGAGGACTTTGCTATG
GATTTGAGGTTGACCTTATGCGCGTAGGTTTTGGTTTTTTTTTGCAGTTTTGGTTTCTTTTGCGGTTTTCTAACC
AATTGCACAACTCCGTTCTCGGGGTGGCGGCAGGCAGGGAGGCTTGGACGCCGGTGGGGAATGGGGGGCCACAG
CTGCAGACCTAAGCCCTCCCCCACCCCTGGAAAGGTCCCTCCCCAACCCAGGCCCCTGGCGTGTGTGGGTGTGCG
TGCGTGTGCGTGCCGTGTTCGTGTGCAAGGGGCCGGGGAGGTGGGCGTGTGTGTGCGTGCCAGCGAAGGCTGCTG
TGGGCGTGTGTGTCAAGTGGGCCACGCGTGCAGGGTGTGTGTCCACGAGCGACGATCGTGGTGGCCCCAGCGGCC
TGGGCGTTGGCTGAGCCGACGCTGGGGCTTCCAGAAGGCCCGGGGGTCTCCGAGGTGCCGGTTAGGAGTTTGAAC
CCCCCCCACTCTGCAGAGGGAAGCGGGGACAATGCCGGGGTTTCAGGCAGGAGACACGAGGAGGGCCTGCCCGGA
AGTCACATCGGCAGCAGCTGTCTAAAGGGCTTGGGGGCCTGGGGGGCGGCGAAGGTGGGTGGGGCCCCTCTGTAA
ATACGGCCCCAGGGTGGTGAGAGAGTCCCATGCCACCCGTCCCCTTGTGACCTCCCCCCTATGACCTCCAGCTGA
CCATGCATGCCACGTGGCTGGCTGGGTCCTCTGCCCTCTTTGGAGTTTGCCTCCCCCAGCCCCCTCCCCATCAAT
AAAACTCTGTTTACAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 544

MQTPRASPPRPALLLLLLLLGGAHGLFPEEPPPLSVAPRDYLNHYPVFVGSGPGRLTPAEGAD
DLNIQRVLRVNRTLFIGDRDNLYRVELEPPTSTELRYQRKLTWRSNPSDINVCRMKGKQEGEC
RNFVKVLLLRDESTLFVCGSNAFNPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALFSDG
MLFTATVTDFLAIDAVIYRSLGDRPTLRTVKHDSKWFKEPYFVHAVEWGSHVYFFFREIAMEF
NYLEKVVVSRVARVCKNDVGGSPRVLEKQWTSFLKARLNCSVPGDSHFYFNVLQAVTGVVSLG
GRPVVLAVFSTPSNSIPGSAVCAFDLTQVAAVFEGRFREQKSPESIWTPVPEDQVPRPRPGCC
AAPGMQYNASSALPDDILNFVKTHPLMDEAVPSLGHAPWILRTLMRHQLTRVAVDVGAGPWGN
QTVVFLGSEAGTVLKFLVRPNASTSGTSGLSVFLEEFETYRPDRCGRPGGGETGQRLLSLELD
AASGGLLAAFPRCVVRVPVARCQQYSGCMKNCIGSQDPYCGWAPDGSCIFLSPGTRAAFEQDV
SGASTSGLGDCTGLLRASLSEDRAGLVSVNLLVTSSVAAFVVGAVVSGFSVGWFVGLRERREL
ARRKDKEAILAHGAGEAVLSVSRLGERRAQGPGGRGGGGGGGAGVPPEALLAPLMQNGWAKAT
LLQGGPHDLDSGLLPTPEQTPLPQKRLPTPHPHPHALGPRAWDHGHPLLPASASSSLLLLAPA
RAPEQPPAPGEPTPDGRLYAARPGRASHGDFPLTPHASPDRRRVVSAPTGPLDPASAADGLPR
PWSPPPTGSLRRPLGPHAPPAATLRRTHTFNSGEARPGDRHRGCHARPGTDLAHLLPYGGADR
TAPPVP

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 318-339, 598-617

N-glycosylation sites.

amino acids 74-78, 155-159, 167-171, 291-295, 386-390, 441-445, 462-466

Glycosaminoglycan attachment sites.

amino acids 51-55, 573-577

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation sites.

amino acids 21-27, 50-56, 189-195, 333-339, 382-388, 448-454, 490-496, 491-497, 508-514, 509-515, 531-537, 558-564, 569-575, 574-580, 580-586, 610-616, 643-649, 663-669, 666-672, 667-673, 668-674, 669-675, 670-676, 868-874, 879-885

FIGURE 545

GATGGCGCAGCCACAGCTTCTGTGAGATTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGG

GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTAT**ATG**CGTCAATTCCCCAAA

ACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT

CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG

ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG

GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGT

CTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG

ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT

GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT

GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC

TATCTGGACAGGGTATTTAAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC

AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCCACATGACA

TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG

GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAG

GAGACAGAA**TAG**GAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT

TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT

GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGATTGTCTTTATGCATCCCC

AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT

TATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTTATTTTTTTACTTGGACATG

AAACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT

ACAGTAAAAAAAAAAAACCTTGTAAATTCTAGAAGAGTGGCTAGGGGGGTTATTCATTTGTAT

TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC

TACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG

ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG

CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAA

FIGURE 546

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPE

IFSSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQE

IRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHY

TLRKISSLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDI

LLQWMEETE

Important features:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 547

AGCAACTCAAGTTCATCATTGTCCTGAGAGAGAGGAGCAGCGCGGTTCTCGGCCGGGACAGCA

GAACGCCAGGGGACCCTCACCTGGGCGCGCCGGGGCACGGGCTTTGATTGTCCTGGGGTCGCG

GAGACCCGCGCGCCTGCCCTGCACGCCGGGCGGCAACCTTTGCAGTCGCGTTGGCTGCTGCGA

TCGGCCGGCGGGTCCCTGCCGAAGGCTCGGCTGCTTCTGTCCACCTCTTACACTTCTTCATTT

ATCGGTGGATCATTTCGAGAGTCCGTCTTGTAA**ATG**TTTGGCACTTTGCTACTTTATTGCTTC

TTTCTGGCGACAGTTCCAGCACTCGCCGAGACCGGCGGAGAAAGGCAGCTGAGCCCGGAGAAG

AGCGAAATATGGGGACCCGGGCTAAAAGCAGACGTCGTCCTTCCCGCCCGCTATTTCTATATT

CAGGCAGTGGATACATCAGGGAATAAATTCACATCTTCTCCAGGCGAAAAGGTCTTCCAGGTG

AAAGTCTCAGCACCAGAGGAGCAATTCACTAGAGTTGGAGTCCAGGTTTTAGACCGAAAAGAT

GGGTCCTTCATAGTAAGATACAGAATGTATGCAAGCTACAAAAATCTGAAGGTGGAAATTAAA

TTCCAAGGGCAACATGTGGCCAAATCCCCATATATTTTAAAAGGGCCGGTTTACCATGAGAAC

TGTGACTGTCCTCTGCAAGATAGTGCAGCCTGGCTACGGGAGATGAACTGCCCTGAAACCATT

GCTCAGATTCAGAGAGATCTGGCACATTTCCCTGCTGTGGATCCAGAAAAGATTGCAGTAGAA

ATCCCAAAAAGATTTGGACAGAGGCAGAGCCTATGTCACTACACCTTAAAGGATAACAAGGTT

TATATCAAGACTCATGGTGAACATGTAGGTTTTAGAATTTTCATGGATGCCATACTACTTTCT

TTGACTAGAAAGGTGAAGATGCCAGATGTGGAGCTCTTTGTTAATTTGGGAGACTGGCCTTTG

GAAAAAAAGAAATCCAATTCAAACATCCATCCGATCTTTTCCTGGTGTGGCTCCACAGATTCC

AAGGATATCGTGATGCCTACGTACGATTTGACTGATTCTGTTCTGGAAACCATGGGCCGGGTA

AGTCTGGATATGATGTCCGTGCAAGCTAACACGGGTCCTCCCTGGGAAAGCAAAAATTCCACT

GCCGTCTGGAGAGGGCGAGACAGCCGCAAAGAGAGACTCGAGCTGGTTAAACTCAGTAGAAAA

CACCCAGAACTCATAGACGCTGCTTTCACCAACTTTTTCTTCTTTAAACACGATGAAAACCTG

TATGGTCCCATTGTGAAACATATTTCATTTTTTGATTTCTTCAAGCATAAGTATCAAATAAAT

ATCGATGGCACTGTAGCAGCTTATCGCCTGCCATATTTGCTAGTTGGTGACAGTGTTGTGCTG

AAGCAGGATTCCATCTACTATGAACATTTTTACAATGAGCTGCAGCCCTGGAAACACTACATT

CCAGTTAAGAGCAACCTGAGCGATCTGCTAGAAAAACTTAAATGGGCGAAAGATCACGATGAA

GAGGCCAAAAAGATAGCAAAAGCAGGACAAGAATTTGCAAGAAATAATCTCATGGGCGATGAC

ATATTCTGTTATTATTTCAAACTTTTCCAGGAATATGCCAATTTACAAGTGAGTGAGCCCCAA

ATCCGAGAGGGCATGAAAAGGGTAGAACCACAGACTGAGGACGACCTCTTCCCTTGTACTTGC

CATAGGAAAAAGACCAAAGATGAACTC**TGA**TATGCAAAATAACTTCTATTAGAATAATGGTGC

TCTGAAGACTCTTCTTAACTAAAAGAAGAATTTTTTTAAGTATTAATTCCATGGACAATATA

AAATCTGTGTGATTGTTTGCAGTATGAAGACACATTTCTACTTATGCAGTATTCTCATGACTG

TACTTTAAAGTACATTTTTAGAATTTTTATAATAAAACCACCTTTATTTTAAAGGAAAAAAA

FIGURE 548

MFGTLLLYCFFLATVPALAETGGERQLSPEKSEIWGPGLKADVVLPARYFYIQAVDTSGNKFT

SSPGEKVFQVKVSAPEEQFTRVGVQVLDRKDGSFIVRYRMYASYKNLKVEIKFQGQHVAKSPY

ILKGPVYHENCDCPLQDSAAWLREMNCPETIAQIQRDLAHFPAVDPEKIAVEIPKRFGQRQSL

CHYTLKDNKVYIKTHGEHVGFRIFMDAILLSLTRKVKMPDVELFVNLGDWPLEKKKSNSNIHP

IFSWCGSTDSKDIVMPTYDLTDSVLETMGRVSLDMMSVQANTGPPWESKNSTAVWRGRDSRKE

RLELVKLSRKHPELIDAAFTNFFFFKHDENLYGPIVKHISFFDFFKHKYQINIDGTVAAYRLP

YLLVGDSVVLKQDSIYYEHFYNELQPWKHYIPVKSNLSDLLEKLKWAKDHDEEAKKIAKAGQE

FARNNLMGDDIFCYYFKLFQEYANLQVSEPQIREGMKRVEPQTEDDLFPCTCHRKKTKDEL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation sites.

amino acids 302-306, 414-418

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 243-247, 495-499

Tyrosine kinase phosphorylation site.

amino acids 341-348

N-myristoylation sites.

amino acids 59-65, 118-124, 184-190, 258-264, 370-376, 439-445

Endoplasmic reticulum targeting sequence.

amino acids 499-504

FIGURE 549

GGGTGATTGAACTAAACCTTCGCCGCACCGAGTTTGCAGTACGGCCGTCACCCGCACCGCTGC
CTGCTTGCGGTTGGAGAAATCAAGGCCCTACCGGGCCTCCGTAGTCACCTCTCTATAGTGGGC
GTGGCCGAGGCCGGGGTGACCCTGCCGGAGCCTCCGCTGCCAGCGAC**ATG**TTCAAGGTAATTC
AGAGGTCCGTGGGGCCAGCCAGCCTGAGCTTGCTCACCTTCAAAGTCTATGCAGCACCAAAAA
AGGACTCACCTCCCAAAAATTCCGTGAAGGTTGATGAGCTTTCACTCTACTCAGTTCCTGAGG
GTCAATCGAAGTATGTGGAGGAGGCAAGGAGCCAGCTTGAAGAAAGCATCTCACAGCTCCGAC
ACTATTGCGAGCCATACACAACCTGGTGTCAGGAAACGTACTCCCAAACTAAGCCCAAGATGC
AAAGTTTGGTTCAATGGGGGTTAGACAGCTATGACTATCTCCAAAATGCACCTCCTGGATTTT
TTCCGAGACTTGGTGTTATTGGTTTTGCTGGCCTTATTGGACTCCTTTTGGCTAGAGGTTCAA
AAATAAAGAAGCTAGTGTATCCGCCTGGTTTCATGGGATTAGCTGCCTCCCTCTATTATCCAC
AACAAGCCATCGTGTTTGCCCAGGTCAGTGGGGAGAGATTATATGACTGGGGTTTACGAGGAT
ATATAGTCATAGAAGATTTGTGGAAGGAGAACTTTCAAAAGCCAGGAAATGTGAAGAATTCAC
CTGGAACTAAG**TAG**AAAACTCCATGCTCTGCCATCTTAATCAGTTATAGGTAAACATTGGAAA
CTCCATAGAATAAATCAGTATTTCTACAGAAAAATGGCATAGAAGTCAGTATTGAATGTATTA
AATTGGCTTTCTTCTTCAGGAAAAACTAGACCAGACCTCTGTTATCTTCTGTGAAATCATCCT
ACAAGCAAACTAACCTGGAATCCCTTCACCTAGAGATAATGTACAAGCCTTAGAACTCCTCAT
TCTCATGTTGCTATTTATGTACCTAATTAAAACCCAAGTTTAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 550

MFKVIQRSVGPASLSLLTFKVYAAPKKDSPPKNSVKVDELSLYSVPEGQSKYVEEARSQLEES
ISQLRHYCEPYTTWCQETYSQTKPKMQSLVQWGLDSYDYLQNAPPGFFPRLGVIGFAGLIGLL
LARGSKIKKLVYPPGFMGLAASLYYPQQAIVFAQVSGERLYDWGLRGYIVIEDLWKENFQKPG
NVKNSPGTK

Important features:
Signal peptide:
Amino acids 1-23

Transmembrane domain:
Amino acids 111-130

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 26-30

Tyrosine kinase phosphorylation site:
Amino acids 36-44

N-myristoylation sites:
Amino acids 124-130;144-150;189-195